华章科技

畅销书全新升级，第1版广获好评；资深MySQL专家撰写，全球知名MySQL数据库服务提供商Percona公司CTO作序推荐，国内多位数据库专家联袂推荐

基于MySQL 5.6，结合源代码，从存储引擎内核角度对InnoDB的整体架构、核心实现和工作机制进行深入剖析

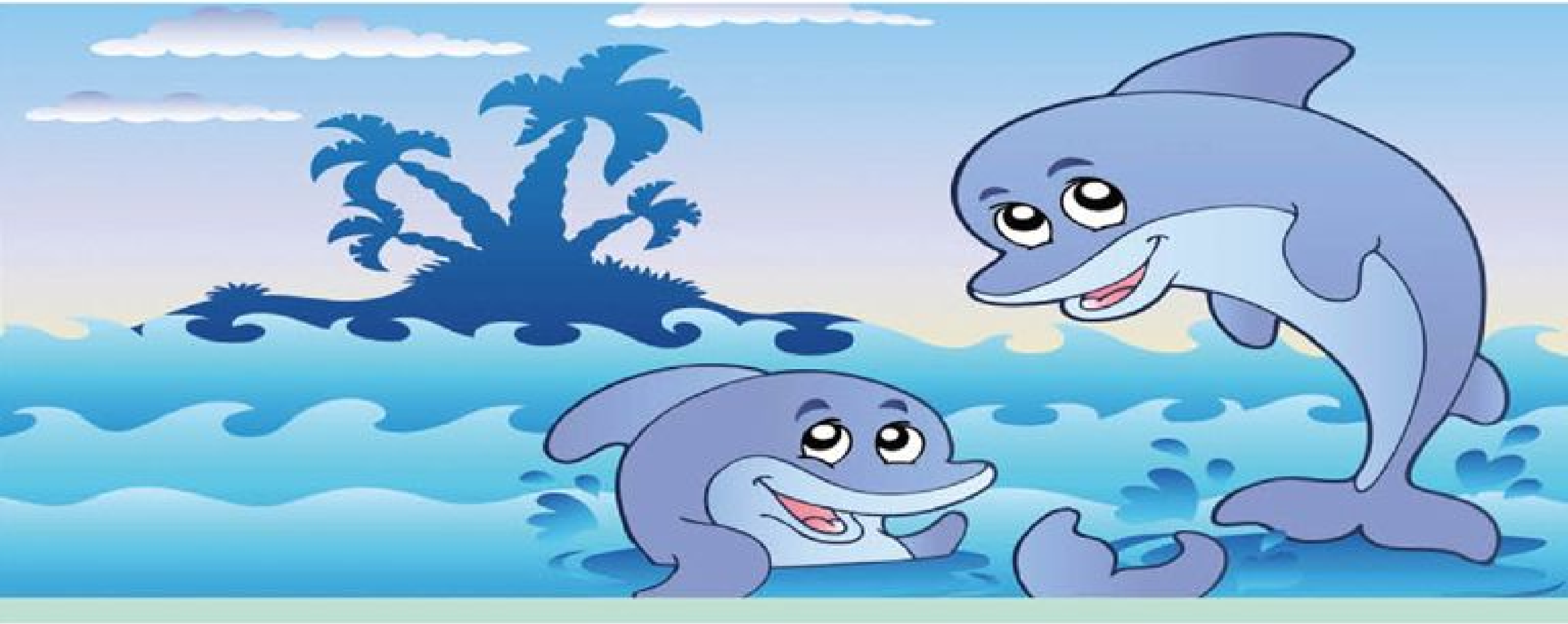
数据库技术丛书

Inside MySQL: InnoDB Storage Engine, Second Edition

MySQL技术内幕

InnoDB存储引擎

第2版



姜承尧◎著

机械工业出版社
China Machine Press

[illegible]

MySQL[illegible]

——InnoDB[illegible] [illegible]2[illegible]

[illegible]　[illegible]

ISBN[illegible]978-7-111-42206-8

[illegible]2013[illegible]

[illegible]

[illegible]+ 86-10-68995265

[illegible]service@bbbvip.com

[illegible]www.hzmedia.com.cn

[illegible] @[illegible]

[illegible] @yanfabook

[illegible]

□□□

It's fair to say that MySQL is the most popular open source database.It has a very large installed base and number of users.Let's see what are the reasons MySQL is so popular,where it stands currently,and maybe touch on some of its future(although predicting the future is rarely successful).

Looking at the customer area of MySQL,which includes Facebook,Flickr,Adobe(in Creative Suite 3),Drupal,Digg,LinkedIn,Wikipedia,eBay,YouTube,Google AdSense(source http://mysql.com/customers/and public resources),it's obvious that MySQL is everywhere.When you log in to your popular forum(powered by Bulleting)or blog(powered by WordPress),most likely it has MySQL as its backend database.Traditionally,two MySQL's characteristics,simplicity of use and performance,were what allowed it to gain such popularity.In addition to that,availability on a very wide range of platforms(including Windows)and built-in replication,which provides an easy scale-out solution for read-only clients,gave more user attractions and production deployments.There is simple evidence of MySQL's simplicity:In 15 minutes or less,you really can get installed,have a working database,and start running queries and store data.From its early stages MySQL had a good interface to most popular languages for Web development-PHP and Perl,and also Java and ODBC connectors.

There are two best known storage engines in MySQL:MyISAM and InnoDB(I don't cover NDB cluster here;it's a totally different story).MyISAM comes as the default storage engine and historically it is the oldest,but InnoDB is ACID compliant and provides transactions,row-level locking,MVCC,automatic recovery and data corruption detection.This makes it the storage engine you want to choose for your application.Also,there is the third-party transaction storage engine PBXT,with characteristics similar to InnoDB,which is included in the MariaDB distribution.

MySQL's simplicity has its own drawback.Just as it is very easy to start working with it,it is very easy to start getting into trouble with it.As soon as your website or forum gets popular,you may figure out that the database is a bottleneck,and that you need special skills and tools to fix it.

The author of this book is a MySQL expert,especially in InnoDB storage engineB.Hence,I highly recommend this book to new users of InnoDB as well as uers who already have well-tuned InnoDB-based applications but need to get internal out of them.

Vadim Tkachenko

□□□□MySQL□□□□□□□□Percona□□CTO

□□MySQL□□□□□MySQLPerformanceBlog.com□□

□□□□MySQL□□2□□□□□□□

[illegible]

[illegible]

[illegible]MySQL[illegible]

MySQL[illegible]InnoDB[illegible]MySQL[illegible]InnoDB[illegible]InnoDB[illegible]InnoDB[illegible]

[illegible]InnoDB[illegible]Think Different[illegible]Think Different[illegible]20[illegible]90[illegible]Think Differently[illegible]Think Different[illegible]Different[illegible]

[illegible]DBA[illegible]“[illegible]”[illegible]“[illegible]”[illegible]Why[illegible]What[illegible]

[illegible]

❑[illegible]“[illegible]”[illegible]

❑[illegible]

❑[illegible]

❑[illegible]

[illegible]MySQL[illegible]MySQL[illegible]SQL[illegible]MySQL[illegible]

MySQL[illegible]MyISAM[illegible]InnoDB[illegible]DBA[illegible]MySQL DBA[illegible]MySQL[illegible]InnoDB[illegible]

[illegible]Why[illegible]What[illegible]InnoDB[illegible]InnoDB[illegible]

[illegible]1[illegible]2[illegible]

[illegible]2[illegible]MySQL 5.6[illegible]InnoDB[illegible]

[illegible]1[illegible]2[illegible]50%[illegible]

❑[illegible]MySQL 5.6[illegible]InnoDB[illegible]MySQL 5.6[illegible]InnoDB[illegible]InnoDB[illegible]

❑[illegible]InnoDB[illegible]redo[illegible]undo[illegible]InnoDB[illegible]undo[illegible]InnoSQL[illegible]undo[illegible]DBA[illegible]InnoDB[illegible]

❑[illegible]6[illegible]InnoDB[illegible]next-key locking[illegible]InnoDB[illegible]

❑[illegible]InnoDB[illegible]insert buffer[illegible]insert buffer[illegible]——change buffer[illegible]

[illegible]

[illegible]MySQL[illegible]SQL[illegible]MySQL InnoDB[illegible]MySQL DBA[illegible]SQL[illegible]MySQL[illegible]Replication[illegible]InnoDB[illegible]

[illegible]InnoDB[illegible]MySQL[illegible]“[illegible]”[illegible]

[illegible]

❑[illegible]SQL[illegible]

❑[illegible]MySQL[illegible]

❑[illegible]C[illegible]C++[illegible]Python[illegible]Java[illegible]

❑[illegible]InnoDB[illegible]

[illegible]

[illegible]10[illegible]“[illegible]”[illegible]10[illegible]InnoDB[illegible]3[illegible]

[illegible]InnoDB[illegible]API[illegible]InnoDB[illegible]

[illegible]www.hzbook.com[illegible]

[illegible]

[illegible]InnoDB[illegible]jiangchengyao@gmail.com[illegible]@insidemysql[illegible]!

[illegible]

[illegible]Pecona[illegible]CEO Peter Zaitsev[illegible]CTO Vadim Tkachenko[illegible]InnoDB[illegible]InnoDB[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]1[illegible] MySQL[illegible]

MySQL[illegible]Linux[illegible]Solaris[illegible]FreeBSD[illegible]Mac[illegible]Windows[illegible]MySQL[illegible]MySQL[illegible]

1.1 [illegible]

[illegible]“[illegible]”[illegible]database[illegible]“[illegible]”[illegible]instance[illegible]

❑[illegible]MySQL[illegible]frm[illegible]MYD[illegible]MYI[illegible]ibd[illegible]NDB[illegible]

❑[illegible]MySQL[illegible]

[illegible]MySQL[illegible]

MySQL[illegible]SQL Server[illegible]Oracle[illegible]Oracle[illegible]Windows[illegible]MySQL[illegible]

[illegible]Linux[illegible]MySQL[illegible]ps[illegible]MySQL[illegible]

```
[root@xen-server bin]#./mysqld_safe[illegible]

[root@xen-server bin]#ps-ef|grep mysqld

root 3441 3258 0 10:23 pts/3 00:00:00/bin/sh./mysqld_safe

mysql 3578 3441 0 10:23 pts/3 00:00:00

/usr/local/mysql/libexec/mysqld--basedir=/usr/local/mysql

--datadir=/usr/local/mysql/var--user=mysql

--log-error=/usr/local/mysql/var/xen-server.err

--pid-file=/usr/local/mysql/var/xen-server.pid

--socket=/tmp/mysql.sock--port=3306

root 3616 3258 0 10:27 pts/3 00:00:00 grep mysqld
```

[illegible]3578[illegible]MySQL[illegible]mysqld_safe[illegible]MySQL[illegible]

[illegible]MySQL[illegible]Oracle[illegible]spfile[illegible]Oracle[illegible]MySQL[illegible]MySQL[illegible]MySQL[illegible]

```
[root@xen-server bin]#mysql--help|grep my.cnf

order of preference,my.cnf,$MYSQL_TCP_PORT,

/etc/my.cnf/etc/mysql/my.cnf/usr/local/mysql/etc/my.cnf[illegible]/.my.cnf
```

[illegible]MySQL[illegible]/etc/my.cnf→/etc/mysql/my.cnf→/usr/local/mysql/etc/my.cnf→[illegible]/.my.cnf[illegible]“[illegible]MySQL[illegible]”[illegible]MySQL[illegible]Linux[illegible]/etc/my.cnf[illegible]Windows[illegible].cnf[illegible].ini[illegible]Windows[illegible]mysql--help[illegible]

```
Default options are read from the following files in the given order:

C:\Windows\my.ini C:\Windows\my.cnf C:\my.ini C:\my.cnf C:\Program Files\MySQL\M

\MySQL Server 5.1\my.cnf
```

[illegible]datadir[illegible]Linux[illegible]datadir[illegible]/usr/local/mysql/data[illegible]

```
mysql[illegible]SHOW VARIABLES LIKE'datadir'\G;

***************************1.row***************************
```

```
Variable_name:datadir

Value:/usr/local/mysql/data/

1 row in set(0.00 sec)1 row in set(0.00 sec)

mysql[illegible]system ls-lh/usr/local/mysql/data

total 32K

drwxr-xr-x 2 root mysql 4.0K Aug 6 16:23 bin

drwxr-xr-x 2 root mysql 4.0K Aug 6 16:23 docs

drwxr-xr-x 3 root mysql 4.0K Aug 6 16:04 include

drwxr-xr-x 3 root mysql 4.0K Aug 6 16:04 lib

drwxr-xr-x 2 root mysql 4.0K Aug 6 16:23 libexec

drwxr-xr-x 10 root mysql 4.0K Aug 6 16:23 mysql-test

drwxr-xr-x 5 root mysql 4.0K Aug 6 16:04 share

drwxr-xr-x 5 root mysql 4.0K Aug 6 16:23 sql-bench

lrwxrwxrwx 1 root mysql 16 Aug 6 16:05 data-[illegible]/opt/mysql_data/
```

[illegible]data[illegible]/opt/mysql_data[illegible]/opt/mysql_data[illegible]mysql[illegible]MySQL[illegible]mysql:mysql[illegible]

1.2 MySQL[illegible]

[illegible]DBA[illegible]MySQL[illegible]MySQL[illegible]SQL Server[illegible]Oracle[illegible]DB2[illegible]

❑[illegible]MySQL[illegible]

❑MySQL[illegible]

❑[illegible]1000[illegible]MySQL[illegible]

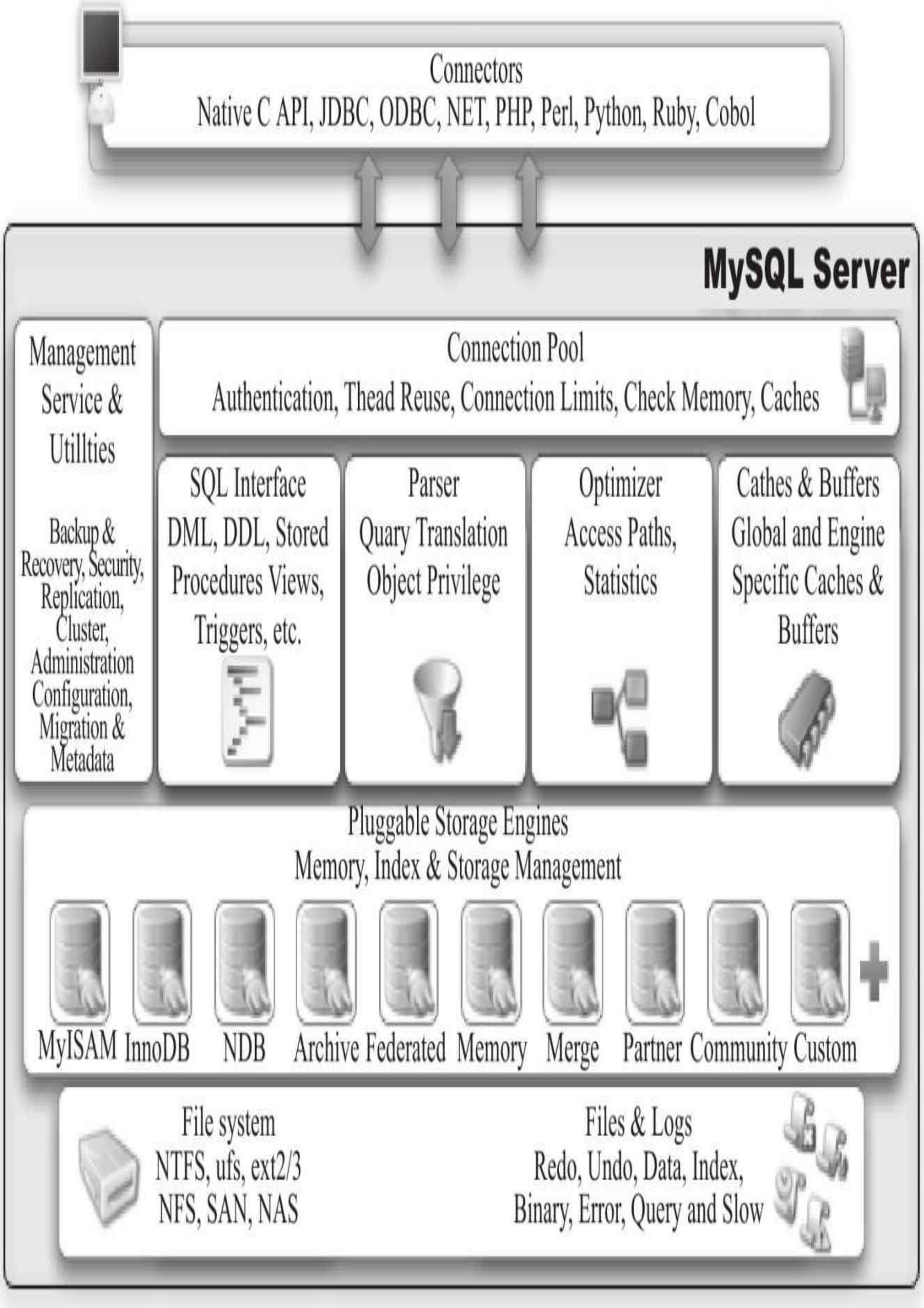
……

[illegible]MySQL[illegible]

[illegible]MySQL[illegible]MySQL[illegible]Oracle[illegible]Microsoft SQL Server[illegible]MySQL[illegible]

[illegible]SELECT[illegible]INSERT[illegible]UPDATE[illegible]DELETE[illegible]Oracle[illegible]SQL Server[illegible]MySQL[illegible]

[illegible]MySQL[illegible]1-1[illegible]MySQL[illegible]


Connectors
Native C API, JDBC, ODBC, NET, PHP, Perl, Python, Ruby, Cobol
MySQL Server
Management Service & Utillties
Backup & Recovery, Security, Replication, Cluster, Administration Configuration, Migration & Metadata
Connection Pool
Authentication, Thead Reuse, Connection Limits, Check Memory, Caches
SQL Interface
DML, DDL, Stored Procedures Views, Triggers, etc.
Parser
Quary Translation Object Privilege
Optimizer
Access Paths, Statistics
Cathes & Buffers
Global and Engine Specific Caches & Buffers
Pluggable Storage Engines
Memory, Index & Storage Management
MyISAM
InnoDB
NDB
Archive
Federated
Memory
Merge
Partner
Community
Custom
File system
NTFS, ufs, ext2/3
NFS, SAN, NAS
Files & Logs
Redo, Undo, Data, Index,
Binary, Error, Query and Slow

[illegible] 1-1 MySQL[illegible]

[illegible]1-1[illegible]MySQL[illegible]

❑[illegible]

❑[illegible]

❑SQL[illegible]

❑[illegible]

❑[illegible]

❑[illegible]Cache[illegible]

❑[illegible]

❑[illegible]

[illegible]1-1[illegible]MySQL[illegible]MySQL[illegible]SQL[illegible]

[illegible]1-1[illegible]MySQL[illegible]MySQL[illegible]

1.3 MySQL□□□□

□□1.2□□□□□□MySQL□□□□□□□□□□□□□□□□□□□□□□□MySQL□□DBA□□□□□□□□□□□□□MySQL□□□□□□□□□□□□□

□□MySQL□□□□□□□□□□□□□□□MySQL□□eBay□□□□Igor Chernyshev□MySQL Memory□□□□□□□□http://code.google.com/p/mysql-heap-dynamic-rows/□□□□□eBay□Personalization Platform□□□□□□□□Google□Facebook□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□SSD□□□□□□[1]□□□□□SSD□□□□□□□□□□□□□□□□□□□□□□□□□□MySQL□□MySQL□□□□□□16□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□MySQL□□□□□□□□□□□□□□□□MySQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□Oracle□□□□□□□□□□□□□□□□MySQL□□□OLTP□Online Transaction Processing□□

1.3.1 InnoDB□□□□

InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□OLTP□□□□□□□□□□□□□□□□□□□□□□□□□Oracle□□□□□□□□□□□□□□□□□□□□MySQL□□□5.5.8□□□□□InnoDB□□□□□□□□□□□□□

InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□MySQL 4.1□□□4.1□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□ibd□□□□□□□InnoDB□□□□□□□□□□□row disk□□□□□□□□□□

InnoDB□□□□□□□□□□□□MVCC□□□□□□□□□□□□□□SQL□□□4□□□□□□□□□REPEATABLE□□□□□□□□□□□□□next-key locking□□□□□□□□

□phantom□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□insert buffer□□□□□□double write□□□□□□□□□□adaptive hash index□□□□□read ahead□□□□□□□□□□□□□

□□□□□□□□□□InnoDB□□□□□□□□□□clustered□□□InnoDB□□□□□□□□□□□□□6□□□ROWID□□□□□□□□□

InnoDB□□□□□MySQL□□□□□□□□□□□□□□Facebook□Google□Yahoo□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□

[1] □ □ □ http://code.google.com/p/david-mysql-tools/wiki/innodb_secondary_buffer_pool

1.3.2　MyISAM□□□□

MyISAM□□□□□□□□□□□□□□□□□□□□□□□□□□□□OLAP□□□□□□□MySQL 5.5.8□□□□MyISAM□□□□□□□□□□□□□□Windows□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□MyISAM□□ETL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□MyISAM□□□□□□□□□□□□□□□□□□□□□□□□□cache□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

MyISAM□□□□□□MYD□MYI□□□MYD□□□□□□□□□MYI□□□□□□□□□□□□□□□myisampack□□□□□□□□□□□□□□□myisampack□□□□□□□□Huffman□□□□□□□□□□□□□□□□□myisampack□□□□□□□□□□□□□□□□□□□□□myisampack□□□□□□□□

□MySQL 5.0□□□□□MyISAM□□□□□□□□□4GB□□□□□□□□□4GB□MyISAM□□□□□□□□MAX_ROWS□AVG_ROW_LENGTH□□□□MySQL 5.0□□□□□MyISAM□□□□256TB□□□□□□□□□□□□□□□□□□

□□　□□MyISAM□□□□□□MySQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□LRU□□□□□□□□□□□□□□□□□□□□□□MySQL 5.1.23□□□□□□□□□32□□□64□□□□□□□□□□□□□□□□□□□□□□□4GB□□□□□□□□□64□□□□□□□□□4GB□□□□□□□

1.3.3 NDB[illegible]

2003[illegible]MySQL AB[illegible]Sony Ericsson[illegible]NDB[illegible]1-1[illegible]NDB[illegible]Oracle[illegible]RAC[illegible]Oracle RAC share everything[illegible]share nothing[illegible]NDB[illegible]MySQL 5.1[illegible]primary key lookups[illegible]NDB[illegible]Data Node[illegible]

[illegible]NDB[illegible]NDB[illegible]JOIN[illegible]MySQL[illegible]NDB[illegible]

[illegible] MySQL NDB Cluster[illegible]NDB Cluster[illegible]Carrier Grade Edition[illegible]http://dev.mysql.com/downloads/cluster/index.html[illegible]NDB Cluster[illegible]

1.3.4 Memory□□□□

Memory□□□□□□□□HEAP□□□Memory□□□□□□□□□□□□□□□□□□□□□B+□□□□

□□Memory□□□TEXT□BLOB□□□□□□□□□□□□□□□□□varchar□□□□□□□□□□char□□□□□□□□□□□□□□□□□□□□□□□□□□□eBay□□□□Igor Chernyshev□□□□□patch□□□□□□

□□□□□□□□□□□MySQL□□□□□Memory□□□□□□□□□□□□□□□□□□□□□intermediate result□□□□□□□□□□□Memory□□□□□□□□□□□□□□□□□□□□□□TEXT□BLOB□□□□□□□MySQL□□□□□□□□□MyISAM□□□□□□□□□□□□□□□□□MyISAM□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

1.3.5 Archive[illegible]

Archive[illegible]INSERT[illegible]SELECT[illegible]MySQL 5.1[illegible]Archive[illegible]zlib[illegible]row[illegible]1:10[illegible]Archive[illegible]Archive[illegible]

1.3.6 Federated[illegible]

Federated[illegible]MySQL[illegible]SQL Server[illegible]Oracle[illegible]Federated[illegible]MySQL[illegible]

1.3.7 Maria[illegible]

Maria[illegible]MyISAM[illegible]MySQL[illegible]Maria[illegible]MySQL[illegible]Michael Widenius[illegible]MyISAM[illegible]Maria[illegible]MVCC[illegible]BLOB[illegible]

1.3.8 [illegible]

[illegible]7[illegible]MySQL[illegible]Merge[illegible]CSV[illegible]Sphinx[illegible]Infobright[illegible]MySQL[illegible]1.2[illegible]

- ❑[illegible]MySQL[illegible]MySQL[illegible]MyISAM[illegible]InnoDB[illegible]1.2[illegible]Sphinx[illegible]

- ❑MySQL[illegible]MySQL[illegible]MyISAM[illegible]InnoDB[illegible]“[illegible]”[illegible]ETL[illegible]MyISAM[illegible]OLTP[illegible]InnoDB[illegible]

- ❑[illegible]1000[illegible]MySQL[illegible]MySQL[illegible]MySQL[illegible]Mytrix[illegible]Inc.[illegible]InnoDB[illegible]1 TB[illegible]InnoDB[illegible]/[illegible]800[illegible]/[illegible]

1.4 [illegible]

[illegible]1.3[illegible]MySQL[illegible]1-2[illegible]MySQL[illegible]MySQL[illegible]MVCC[illegible]

Feature	MyISAM	BDB	Memory	InnoDB	Archive	NDB
Storage Limits	No	No	Yes	64TB	No	Yes
Transactions (commit, rollback, etc.)		✓		✓		
Locking granularity	Table	Page	Table	Row	Row	Row
MVCC/Snapshot Read				✓	✓	✓
Geospatial support	✓					
B-Tree indexes	✓	✓	✓	✓		✓
Hash indexes			✓	✓		✓
Full text search index	✓					
Clustered index				✓		
Data Caches			✓	✓		✓
Index Caches	✓		✓	✓		✓
Compressed data	✓				✓	
Encrypted data (via function)	✓	✓	✓	✓	✓	✓
Storage cost (space used)	Low	Low	N/A	High	Very Low	Low
Memory cost	Low	Low	Medium	High	Low	High
Bulk Insert Speed	High	High	High	Low	Very High	High
Cluster database support						✓
Replication support	✓	✓	✓	✓	✓	✓
Foreign key support				✓		
Backup/Point-in-time recovery	✓	✓	✓	✓	✓	✓
Query cache support	✓	✓	✓	✓	✓	✓
Update Statistics for Data Dictionary	✓	✓	✓	✓	✓	✓

[illegible] 1-2 [illegible]MySQL[illegible]

[illegible]MySQL[illegible]MySQL[illegible]MySQL[illegible]

[illegible]SHOW ENGINES[illegible]MySQL[illegible]information_schema[illegible]ENGINES[illegible]

```
mysql[illegible]SHOW ENGINES\G;

***************************1.row***************************

Engine:InnoDB

Support:YES

Comment:Supports transactions,row-level locking,and foreign keys

Transactions:YES

XA:YES

Savepoints:YES

***************************2.row***************************

Engine:MRG_MYISAM

Support:YES

Comment:Collection of identical MyISAM tables

Transactions:NO

XA:NO

Savepoints:NO

***************************3.row***************************

Engine:BLACKHOLE

Support:YES

Comment:/dev/null storage engine(anything you write to it disappears)

Transactions:NO

XA:NO

Savepoints:NO
```

```
***************************4.row***************************

Engine:CSV

Support:YES

Comment:CSV storage engine

Transactions:NO

XA:NO

Savepoints:NO

***************************5.row***************************

Engine:MEMORY

Support:YES

Comment:Hash based,stored in memory,useful for temporary tables

Transactions:NO

XA:NO

Savepoints:NO

***************************6.row***************************

Engine:FEDERATED

Support:NO

Comment:Federated MySQL storage engine

Transactions:NULL

XA:NULL

Savepoints:NULL

***************************7.row***************************

Engine:ARCHIVE

Support:YES

Comment:Archive storage engine

Transactions:NO

XA:NO

Savepoints:NO

***************************8.row***************************

Engine:MyISAM
```

```
Support:DEFAULT

Comment:Default engine as of MySQL 3.23 with great performance

Transactions:NO

XA:NO

Savepoints:NO

8 rows in set(0.00 sec)
```

[illegible]MySQL[illegible]

```
mysql[illegible]CREATE TABLE mytest Engine=MyISAM

-[illegible]AS SELECT*FROM salaries;

Query OK,2844047 rows affected(4.37 sec)

Records:2844047 Duplicates:0 Warnings:0

mysql[illegible]ALTER TABLE mytest Engine=InnoDB;

Query OK,2844047 rows affected(15.86 sec)

Records:2844047 Duplicates:0 Warnings:0

mysql[illegible]ALTER TABLE mytest Engine=ARCHIVE;

Query OK,2844047 rows affected(16.03 sec)

Records:2844047 Duplicates:0 Warnings:0
```

[illegible]MyISAM[illegible]40.7MB[illegible]InnoDB[illegible]113.6MB[illegible]Archive[illegible]20.2MB[illegible]

[illegible] MySQL[illegible]MySQL[illegible]SQL Server[illegible]AdventureWorks[illegible]Oracle[illegible]MySQL[illegible]Oracle[illegible]SQL Server[illegible]http://dev.mysql.com/doc/[illegible]

1.5 □□MySQL

□□□□□□□MySQL□□□□□□□□□□□□□□□□□□MySQL□□□□□□□□□□MySQL□□TCP/IP□□□□UNIX□□□□□MySQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

1.5.1 TCP/IP

TCP/IP□□□□□□MySQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□TCP/IP□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□client□□□□□□□MySQL□□□server□□□□□□□□□□□□□□□□□□□TCP/IP□□□□□□□□□□□□Windows□□□□□□□□□□Linux□□□□□MySQL□□□□□□□□

```
C:\□mysql-h192.168.0.101-u david-p

Enter password:

Welcome to the MySQL monitor.Commands end with;or\g.

Your MySQL connection id is 18358

Server version:5.0.77-log MySQL Community Server(GPL)

Type'help;'or'\h'for help.Type'\c'to clear the current input statement.

mysql□
```

□□□□□□□Windows□□□□□Host IP□192.168.0.101□MySQL□□□□□TCP/IP□□□□□□□□□□□□□□□□□□MySQL□□□□□□□□□□□□□□DDL□DML□□

□□□□□□□□□□□□TCP/IP□□□MySQL□□□□MySQL□□□□□□□□□□□□□□□□□□□□□□□□□□IP□□□□□□□MySQL□□□□□□□mysql□□□□□□□user□□□□□□□

```
mysql□USE mysql;

Database changed
```

```
mysql[illegible]SELECT host,user,password FROM user;

***************************1.row*****************************

host:192.168.24.%

user:root

password:*75DBD4FA548120B54FE693006C41AA9A16DE8FBE

****************************2.row*****************************

host:nineyou0-43

user:root

password:*75DBD4FA548120B54FE693006C41AA9A16DE8FBE

****************************3.row*****************************

host:127.0.0.1

user:root

password:*75DBD4FA548120B54FE693006C41AA9A16DE8FBE

****************************4.row*****************************

host:192.168.0.100

user:zlm

password:*DAE0939275CC7CD8E0293812A31735DA9CF0953C

****************************5.row*****************************

host:%

user:david

password:

5 rows in set(0.00 sec)
```

[illegible]MySQL[illegible]david[illegible]IP[illegible]root[illegible]

1.5.2 [illegible]

[illegible]Windows 2000[illegible]Windows XP[illegible]Windows 2003[illegible]Windows Vista [illegible] Microsoft SQL Server[illegible]MySQL[illegible]--enable-named-pipe[illegible]MySQL 4.1[illegible] MySQL[illegible]--shared-memory[illegible]MySQL[illegible]--protocol=memory[illegible]

1.5.3 UNIX□□□□

□Linux□UNIX□□□□□□□□□UNIX□□□□□UNIX□□□□□□□□□□□□□□□□□□□□MySQL□□□--socket=/tmp/mysql.sock□□□□□□□□□□□□□□□□□□□□□□□□UNIX□□□□□□□□□□

```
mysql□SHOW VARIABLES LIKE'socket';

***************************1.row***************************

Variable_name:socket

Value:/tmp/mysql.sock

1 row in set(0.00 sec)
```

□□□□UNIX□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
[root@stargazer□]#mysql-udavid-S/tmp/mysql.sock

Welcome to the MySQL monitor.Commands end with;or\g.

Your MySQL connection id is 20333

Server version:5.0.77-log MySQL Community Server(GPL)

Type'help;'or'\h'for help.Type'\c'to clear the buffer.

mysql□
```

1.6 [illegible]

[illegible]MySQL[illegible]“[illegible]”[illegible]“[illegible]”[illegible]MySQL DBA[illegible]MySQL[illegible]MySQL[illegible]MySQL[illegible]MySQL[illegible]

[illegible]MySQL[illegible]

□2□ InnoDB□□□□

InnoDB□□□□□□MySQL□□□□□□□□□□□□□□Oracle□□□□□□□□□□□□InnoDB□□□□□OLTP□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□MySQL□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□

2.1 InnoDB□□□□□□

InnoDB□□□□□□□Innobase Oy□□[1]□□□□□□□MySQL□□□□□□□□□□□□□□□□MySQL 5.5□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□Windows□□□□□□□□□□□□□□□□□□□□□□□□ACID□□□MySQL□□□□□BDB□□□□□□□□□MySQL□□□□□□□□□□□□□□□□□□□□□□□□□□MVCC□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□CPU□

Heikki Tuuri□1964□□□□□□□□□□InnoDB□□□□□□□□□□□□□Linux□□□Linus□□□□□□□□□□□□□1990□□□□□□□□□□□□□□□□□□□□□□1995□□□Innobase Oy□□□□□CEO□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□Calvin Sun□□□□□□□□□□□Jimmy Yang□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□Sybase□□□□□□□□□□□□□□□□□□□

InnoDB□□□□□□□□□□□□□□□□□□□□□□Google□Yahoo!□Facebook□YouTube□Flickr□□□□□□□□□□□□□□□□□Second Life□□□□□□□□□□□□□MySQL□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□OLTP□□□□□MySQL InnoDB□□□□□□□□□□□□

□MySQL□□□□□□□□□□□□□□□Internet□□□□Slashdot.org□□□InnoDB□□Mytrix□Inc.□InnoDB□□□□□1 TB□□□□□□□□□□□□□InnoDB□□□□□/□□□□□□□□□□800□/□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□

InnoDB□□□□□MySQL□□□□□□□GNU GPL 2□□□□□□□□MySQL□□□□□□□□□http://www.mysql.com/about/legal/□□□□□□□□□□

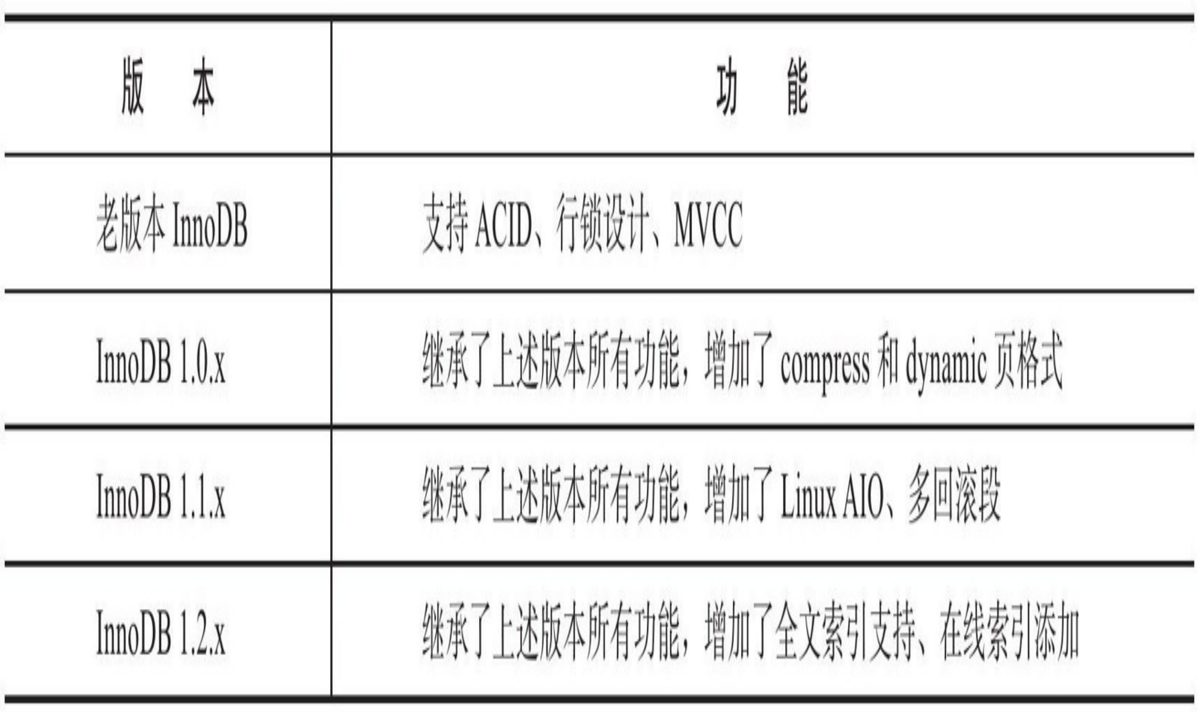
---

[1] 2006□□□□□□□Oracle□□□□□

2.2 InnoDB□□□□□□□

InnoDB□□□□□□□□□□MySQL□□□□□□□□□□□□□□□□□□□□MySQL□□□□□□□□□□□MySQL 5.1□□□□MySQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□MySQL□□□□□□□□□□□□MySQL 5.1□□□□□□□□□□□InnoDB□□□□□□□□□InnoDB□□□□□□□□□□□□InnoDB□□□□□□□□□□InnoDB□□□□□□□InnoDB Plugin□□□□□□InnoDB 1.0.x□□□MySQL 5.5□□□□□InnoDB□□□□□□□1.1.x□□□□□□MySQL 5.6□□□InnoDB□□□□□□□□□1.2.x□□□□2-1□□□□□□□□InnoDB□□□□□□□□

表2-1 InnoDB各版本功能对比



版　本	功　能
老版本 InnoDB	支持 ACID、行锁设计、MVCC
InnoDB 1.0.x	继承了上述版本所有功能，增加了 compress 和 dynamic 页格式
InnoDB 1.1.x	继承了上述版本所有功能，增加了 Linux AIO、多回滚段
InnoDB 1.2.x	继承了上述版本所有功能，增加了全文索引支持、在线索引添加

□□□□□□□□□□□MySQL□□□□□□□□MySQL 5.1□□□□□□InnoDB Plugin□□□DBA□□□□□InnoDB Plugin□InnoDB 1.1□□□□□□□□□□□□□□2-1□□□□□□□□□□□□□□□□compress□dynamic□□□□□□□InnoDB Plugin□□□□Linux Native AIO□□□□□□□□□□□□□□□□□InnoDB

Plugin[illegible]1023[illegible]MySQL 5.5[illegible]
InnoDB Plugin[illegible]

2.3 InnoDB[illegible]

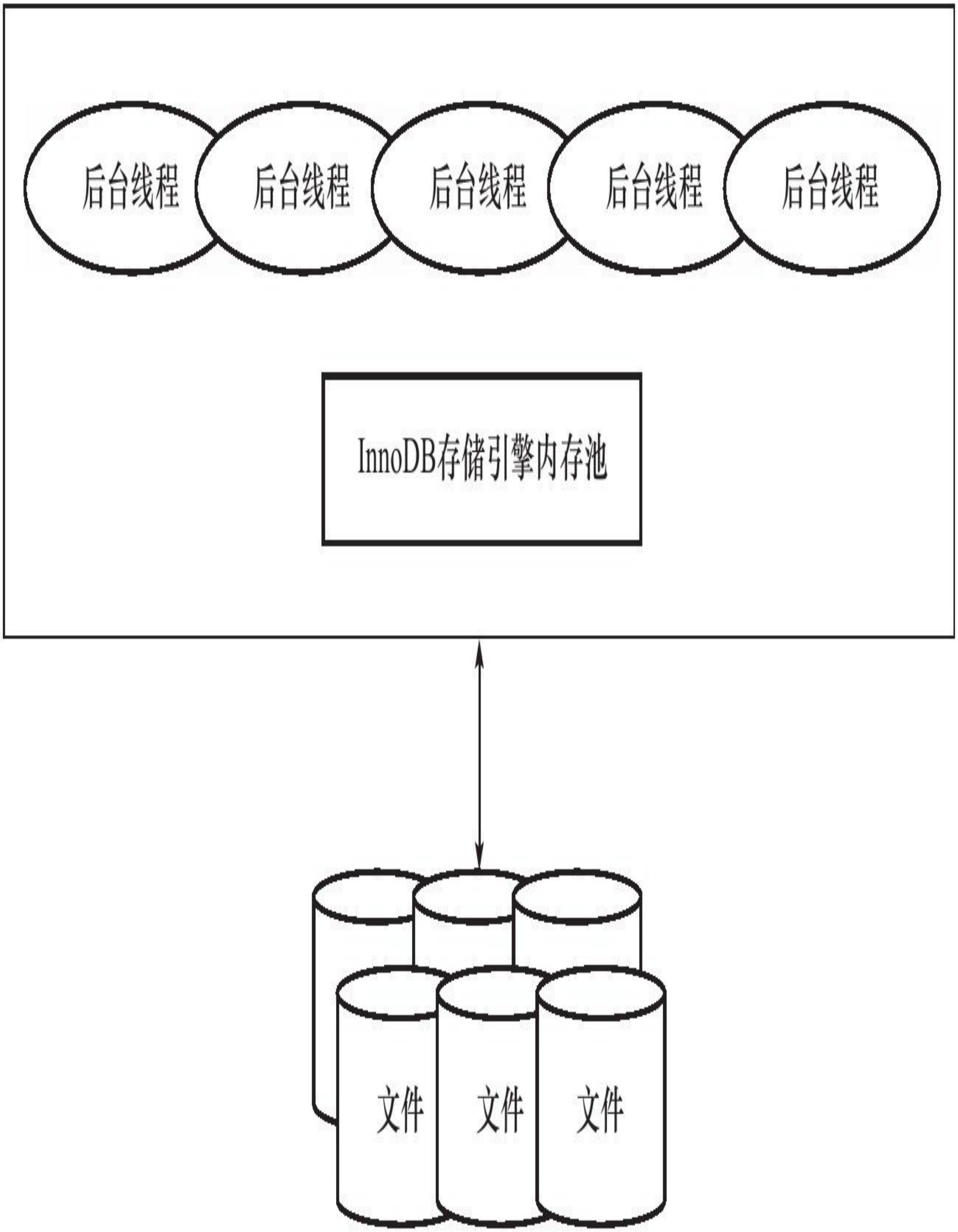
[illegible]1[illegible]MySQL[illegible]InnoDB[illegible]2-1[illegible]InnoDB[illegible]InnoDB[illegible]

❑[illegible]/[illegible]

❑[illegible]

❑[illegible]redo log[illegible]

……



[illegible] 2-1 InnoDB[illegible]

□□□InnoDB□□□□□□□□□□□

2.3.1 □□□□

InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

1.Master Thread

Master Thread□□□INSERT BUFFER□□UNDO□□□□□□2.5□□□□□□□□□□□□Master Thread□□□□□□

2.IO Thread

□InnoDB□□□□□□□□□□AIO□Async IO□□□□□IO□□□□□□□□□□□□□□□□□□□IO Thread□□□□□□□□□□IO□□□□□□call back□□□□InnoDB 1.0□□□□□□4□IO Thread□□□□write□read□insert buffer□log IO thread□□Linux□□□□IO Thread□□□□□□□□□□□□□Windows□□□□□□□□□innodb_file_io_threads□□□IO Thread□□InnoDB 1.0.x□□□□□read thread□write thread□□□□□□4□□□□□□□□innodb_file_io_threads□□□□□□□□□innodb_read_io_threads□innodb_write_io_threads□□□□□□□□□

```
mysql□SHOW VARIABLES LIKE'innodb_version'\G;

***************************1.row***************************

Variable_name:innodb_version

Value:1.0.6

1 row in set(0.00 sec)

mysql□SHOW VARIABLES LIKE'innodb_%io_threads'\G;

***************************1.row***************************

Variable_name:innodb_read_io_threads

Value:4

***************************2.row***************************
```

```
Variable_name:innodb_write_io_threads
Value:4
2 rows in set(0.00 sec)
```

□□□□□SHOW ENGINE INNODB STATUS□□□InnoDB□□IO Thread□

```
mysql□SHOW ENGINE INNODB STATUS\G;
***************************1.row***************************
Type:InnoDB
Name:
Status:
=====================================
100719 21:55:26 INNODB MONITOR OUTPUT
=====================================
Per second averages calculated from the last 36 seconds
……
--------
FILE I/O
--------
I/O thread 0 state:waiting for i/o request(insert buffer thread)
I/O thread 1 state:waiting for i/o request(log thread)
I/O thread 2 state:waiting for i/o request(read thread)
I/O thread 3 state:waiting for i/o request(read thread)
I/O thread 4 state:waiting for i/o request(read thread)
I/O thread 5 state:waiting for i/o request(read thread)
I/O thread 6 state:waiting for i/o request(write thread)
I/O thread 7 state:waiting for i/o request(write thread)
I/O thread 8 state:waiting for i/o request(write thread)
I/O thread 9 state:waiting for i/o request(write thread)
```

```
……
----------------------------
END OF INNODB MONITOR OUTPUT
============================
1 row in set(0.01 sec)
```

□□□□IO Thread 0□insert buffer thread□IO Thread 1□log thread□□□□□□□□□innodb_read_io_threads□innodb_write_io_threads□□□□□□□□□□□□□□□ID□□□□□□□□

3.Purge Thread

□□□□□□□□□□□□undolog□□□□□□□□□□□PurgeThread□□□□□□□□□□□undo□□□InnoDB 1.1□□□□□purge□□□□InnoDB□□□□□Master Thread□□□□□□InnoDB 1.1□□□□□purge□□□□□□□□□□□□□□□□□□□□□Master Thread□□□□□□□□CPU□□□□□□□□□□□□□□□□□□□□□MySQL□□□□□□□□□□□□□□□□□□□□□Purge Thread□

```
[mysqld]
innodb_purge_threads=1
```

□InnoDB 1.1□□□□□□□innodb_purge_threads□□□□1□InnoDB□□□□□□□□□□□□□1□□□□□□□□□□□□□□□□□□

```
120529 22:54:16[Warning]option'innodb-purge-threads':unsigned value 4 adjusted to 1
```

□InnoDB 1.2□□□□□InnoDB□□□□Purge Thread□□□□□□□□□□□□□□□undo□□□□□□□□□Purge Thread□□□□□□□undo□□□□□□□□□□□□□□□□□□□□□□□□□□□□□4□Purge Thread□

```
mysql□SELECT VERSION()\G;
***************************1.row***************************
```

```
VERSION():5.6.6

1 row in set(0.00 sec)

mysql[illegible]SHOW VARIABLES LIKE'innodb_purge_threads'\G;

***************************1.row***************************

Variable_name:innodb_purge_threads

Value:4

1 row in set(0.00 sec)
```

4.Page Cleaner Thread

Page Cleaner Thread[illegible]InnoDB 1.2.x[illegible]Master Thread[illegible]InnoDB[illegible]

2.3.2 [illegible]

1.[illegible]

InnoDB[illegible]Disk-base Database[illegible]CPU[illegible]

[illegible]“FIX”[illegible]

[illegible]Checkpoint[illegible]

[illegible]32[illegible]3G[illegible]PAE[illegible]32[illegible]64GB[illegible]8GB[illegible]PC[illegible]512GB[illegible]64[illegible]

[illegible]InnoDB[illegible]innodb_buffer_pool_size[illegible]MySQL[illegible]InnoDB[illegible]15GB[illegible]

```
mysql□SHOW VARIABLES LIKE'innodb_buffer_pool_size'\G;

***************************1.row***************************

Variable_name:innodb_buffer_pool_size

Value:16106127360

1 row in set(0.00 sec)
```

[illegible]undo[illegible]insert buffer[illegible]adaptive hash index[illegible]InnoDB[illegible]lock info[illegible]

□□□□data dictionary□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ □□□□□□□□□□2-2□□□□□□InnoDB□□□□□□□□□□□□□

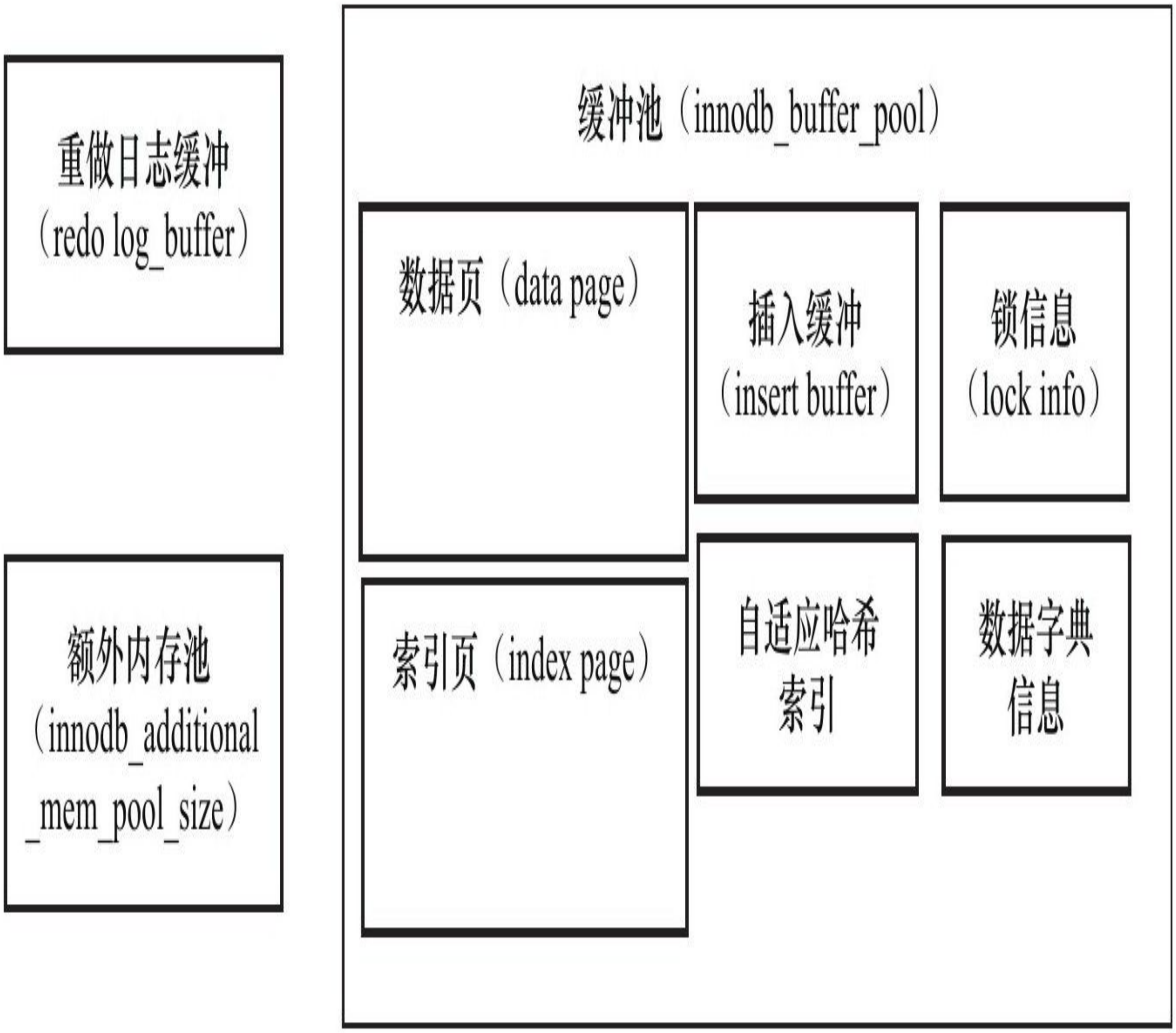


图 2-2 InnoDB□□□□□□

□InnoDB 1.0.x□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ innodb_buffer_pool_instances□□□□□□□□□□□1□

```
mysql＞SHOW VARIABLES LIKE'innodb_buffer_pool_instances'\G;

***************************1.row***************************
```

```
Variable_name:innodb_buffer_pool_instances

Value:1

1 row in set(0.00 sec)
```

□□□□□□□innodb_buffer_pool_instances□□□□□□1□□□□□□□□□□□□□□□□□□□□SHOW ENGINE INNODB STATUS□□□□□□□□□□□

```
mysql□SHOW ENGINE INNODB STATUS\G;

***************************1.row***************************

Type:InnoDB

……

----------------------

INDIVIDUAL BUFFER POOL INFO

----------------------

---BUFFER POOL 0

Buffer pool size 65535

Free buffers 65451

Database pages 84

Old database pages 0

Modified db pages 0

Pending reads 0

Pending writes:LRU 0,flush list 0 single page 0

Pages made young 0,not young 0

0.00 youngs/s,0.00 non-youngs/s

Pages read 84,created 0,written 1

9.33 reads/s,0.00 creates/s,0.11 writes/s

Buffer pool hit rate 764/1000,young-making rate 0/1000 not 0/1000

Pages read ahead 0.00/s,evicted without access 0.00/s,Random read ahead 0.00/s

LRU len:84,unzip_LRU len:0

I/O sum[0]:cur[0],unzip sum[0]:cur[0]
```

```
---BUFFER POOL 1
Buffer pool size 65536
Free buffers 65473
Database pages 63
Old database pages 0
Modified db pages 0
Pending reads 0
Pending writes:LRU 0,flush list 0 single page 0
Pages made young 0,not young 0
0.00 youngs/s,0.00 non-youngs/s
Pages read 63,created 0,written 0
7.00 reads/s,0.00 creates/s,0.00 writes/s
Buffer pool hit rate 500/1000,young-making rate 0/1000 not 0/1000
Pages read ahead 0.00/s,evicted without access 0.00/s,Random read ahead 0.00/s
LRU len:63,unzip_LRU len:0
I/O sum[0]:cur[0],unzip sum[0]:cur[0]
```

□□□□□innodb_buffer_pool_instances□□□2□□□□□□□□□□□□□□□□□□□□□SHOW ENGINE INNODB STATUS□□□□□□□□□□□□□□□□□□□□□□□□□□---BUFFER POOL 0□□□□□□□□□□□□□□□

□MySQL 5.6□□□□□□□□□□information_schema□□□□□INNODB_BUFFER_POOL_STATS□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
mysql□SELECT POOL_ID,POOL_SIZE,
-□FREE_BUFFERS,DATABASE_PAGES
-□FROM INNODB_BUFFER_POOL_STATS\G;
***************************1.row***************************
POOL_ID:0
POOL_SIZE:65535
```

```
FREE_BUFFERS:65451

DATABASE_PAGES:84

***************************2.row***************************

POOL_ID:1

POOL_SIZE:65536

FREE_BUFFERS:65473

DATABASE_PAGES:63
```

2.LRU List□Free List□Flush List

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□LRU□Latest Recent Used□□□□□□□□□□□□□□□□□□□□□□□□□□LRU□□□□□□□□□□□□□□LRU□□□□□□□□□□□□□□□□□□□□□□□□□□□LRU□□□□□□□□

□InnoDB□□□□□□□□□□□□□□□□□16KB□□□□□LRU□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□LRU□□□□□□□□□□InnoDB□□□□□□□LRU□□□□□□□midpoint□□□□□□□□□□□□□□□□□□□□□□□□□□□□□LRU□□□□□□□□□□□LRU□□□midpoint□□□□□□□□InnoDB□□□□□□□midpoint insertion strategy□□□□□□□□□□□□LRU□□□□□5/8□□midpoint□□□□□□innodb_old_blocks_pct□□□□□

```
mysql□SHOW VARIABLES LIKE'innodb_old_blocks_pct'\G;

***************************1.row***************************

Variable_name:innodb_old_blocks_pct

Value:37

1 row in set(0.00 sec)
```

□□□□□□□□□□□□□innodb_old_blocks_pct□□□□37□□□□□□□□□□□LRU□□□□□37%□□□□□□□3/8□□□□□□InnoDB□□□□□□□midpoint□□□□□□□old□□□□□□□□□□new□□□□□□□□□□□new□□□□□□□□□□□□□□□□□

[illegible]LRU[illegible]LRU[illegible]LRU[illegible]SQL[illegible]LRU[illegible]LRU[illegible]InnoDB[illegible]

[illegible]InnoDB[illegible]LRU[illegible]innodb_old_blocks_time[illegible]mid[illegible]LRU[illegible]SQL[illegible]LRU[illegible]

```
mysql＞SET GLOBAL innodb_old_blocks_time=1000;

Query OK,0 rows affected(0.00 sec)

#data or index scan operation

……

mysql＞SET GLOBAL innodb_old_blocks_time=0;

Query OK,0 rows affected(0.00 sec)
```

[illegible]63%[illegible]SQL[illegible]

```
mysql＞SET GLOBAL innodb_old_blocks_pct=20;

Query OK,0 rows affected(0.00 sec)
```

LRU[illegible]LRU[illegible]Free[illegible]Free[illegible]Free[illegible]LRU[illegible]LRU[illegible]LRU[illegible]LRU[illegible]old[illegible]new[illegible]page made young[illegible]innodb_old_blocks_time[illegible]old[illegible]new[illegible]page not made young[illegible]SHOW ENGINE INNODB STATUS[illegible]LRU[illegible]Free[illegible]

```
mysql＞SHOW ENGINE INNODB STATUS\G;
```

```
**************************1.row**************************

Type:InnoDB

Name:

Status:

=====================================

120725 22:04:25 INNODB MONITOR OUTPUT

=====================================

Per second averages calculated from the last 24 seconds

……

Buffer pool size 327679

Free buffers 0

Database pages 307717

Old database pages 113570

Modified db pages 24673

Pending reads 0

Pending writes:LRU 0,flush list 0,single page 0

Pages made young 6448526,not young 0

48.75 youngs/s,0.00 non-youngs/s

Pages read 5354420,created 239625,written 3486063

55.68 reads/s,81.74 creates/s,955.88 writes/s

Buffer pool hit rate 1000/1000,young-making rate 0/1000 not 0/1000

……
```

[illegible]SHOW ENGINE INNODB STATUS[illegible]Buffer pool size[illegible]327 679[illegible]327679*16K[illegible]5GB[illegible]Free buffers[illegible]Free[illegible]Database pages[illegible]LRU[illegible]Free buffers[illegible]Database pages[illegible]Buffer pool size[illegible]2-2[illegible]Lock[illegible]Insert Buffer[illegible]LRU[illegible]LRU[illegible] pages made young□□□LRU□□□□□□□□□□□□□□□□□□□□□□□□□□□□innodb_old_blocks_time□□□□□not young□0□youngs/s□non-youngs/s□□□□□□□□□□□□□□□□□□□□□□□□□□——Buffer pool hit rate□□□□□□□□□□□□□□□□□100%□□□□□□□□□□□□□□□□□□□□□□□□95%□□□□Buffer pool hit rate□□□□95%□□□□□□□□□□□□□□□□□□□□□□□□LRU□□□□□□□□□

□□ □□□□SHOW ENGINE INNODB STATUS□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□Per second averages calculated from the last 24 seconds□□□□□□□□24□□□□□□□□□

□InnoDB 1.2□□□□□□□□□□□INNODB_BUFFER_POOL_STATS□□□□□□□□□□□□□□

```
mysql□SELECT POOL_ID,HIT_RATE,

-□PAGES_MADE_YOUNG,PAGES_NOT_MADE_YOUNG

-□FROM information_schema.INNODB_BUFFER_POOL_STATS\G;

***************************1.row***************************

POOL_ID:0

HIT_RATE:980

PAGES_MADE_YOUNG:450

PAGES_NOT_MADE_YOUNG:0
```

□□□□□□□□□INNODB_BUFFER_PAGE_LRU□□□□□LRU□□□□□□□□□□□□□□□□□□□□□□□□□□□LRU□□□SPACE□1□□□□□□□

```
mysql□SELECT TABLE_NAME,SPACE,PAGE_NUMBER,PAGE_TYPE

-□FROM INNODB_BUFFER_PAGE_LRU WHERE SPACE=1;

+------------+-------+-------------+------------------+

|TABLE_NAME|SPACE|PAGE_NUMBER|PAGE_TYPE|

+------------+-------+-------------+------------------+

|NULL|1|0|FILE_SPACE_HEADER|
|NULL|1|1|IBUF_BITMAP|

|NULL|1|2|INODE|

|test/t|1|3|INDEX|

+------------+-------+------------+------------------+

4 rows in set(0.00 sec)
```

InnoDB[illegible]1.0.x[illegible]16KB[illegible]1KB[illegible]2KB[illegible]4KB[illegible]8KB[illegible]LRU[illegible]16KB[illegible]unzip_LRU[illegible]SHOW ENGINE INNODB STATUS[illegible]

```
mysql[illegible]SHOW ENGINE INNODB STATUS\G;

……

Buffer pool hit rate 999/1000,young-making rate 0/1000 not 0/1000

Pages read ahead 0.00/s,evicted without access 0.00/s,Random read ahead 0.00/s

LRU len:1539,unzip_LRU len:156

I/O sum[0]:cur[0],unzip sum[0]:cur[0]

……
```

[illegible]LRU[illegible]1539[illegible]unzip_LRU[illegible]156[illegible]LRU[illegible]unzip_LRU[illegible]

[illegible]8KB[illegible]2KB[illegible]unzip_LRU[illegible]

[illegible]unzip_LRU[illegible]4KB[illegible]

1[illegible]4KB[illegible]unzip_LRU[illegible]

2[illegible]

3[illegible]8KB[illegible]unzip_LRU[illegible]

4）若能够得到空闲页，将页分成2个4KB页，存放到4KB的unzip_LRU列表；

5）若不能得到空闲页，从LRU列表中申请一个16KB的页，将页分为1个8KB的页、2个4KB的页，分别存放到对应的unzip_LRU列表中。

同样可以通过information_schema架构下的表INNODB_BUFFER_PAGE_LRU来观察unzip_LRU列表中的页，如：

```
mysql>SELECT
->TABLE_NAME,SPACE,PAGE_NUMBER,COMPRESSED_SIZE
->FROM INNODB_BUFFER_PAGE_LRU
->WHERE COMPRESSED_SIZE<>0;
+------------+-------+------------+----------------+
|TABLE_NAME|SPACE|PAGE_NUMBER|COMPRESSED_SIZE|
+------------+-------+------------+----------------+
|sbtest/t|9|134|8192|
|sbtest/t|9|135|8192|
|sbtest/t|9|96|8192|
|sbtest/t|9|136|8192|
|sbtest/t|9|32|8192|
|sbtest/t|9|97|8192|
|sbtest/t|9|137|8192|
|sbtest/t|9|98|8192|
……
```

在LRU列表中的页被修改后，称该页为脏页（dirty page），即缓冲池中的页和磁盘上的页的数据产生了不一致。这时数据库会通过CHECKPOINT机制将脏页刷新回磁盘，而Flush列表中的页即为脏页列表。需要注意的是，脏页既存在于LRU列表中，也存在于Flush列表中。LRU列表用来管理缓冲池中页的可用性，Flush列表用来管理将页刷新回磁盘，二者互不影响。

同LRU列表一样，Flush列表也可以通过命令SHOW ENGINE INNODB STATUS来查看，前面例子中Modified db pages 24673就显示了脏页的数量。

information_schema[illegible]INNODB_BUFFER_PAGE_LRU[illegible]LRU[illegible]INNODB_BUFFER_PAGE_LRU[illegible]OLDEST_MODIFICATION[illegible]0[illegible]SQL[illegible]

```
mysql[illegible]SELECT TABLE_NAME,SPACE,PAGE_NUMBER,PAGE_TYPE

-[illegible]FROM INNODB_BUFFER_PAGE_LRU

-[illegible]WHERE OLDEST_MODIFICATION[illegible]0;

+------------+-------+-------------+-------------------+

|TABLE_NAME|SPACE|PAGE_NUMBER|PAGE_TYPE|

+------------+-------+-------------+-------------------+

|NULL|0|56|SYSTEM|

|NULL|0|0|FILE_SPACE_HEADER|

|test/t|1|3|INDEX|

|NULL|0|320|INODE|

|NULL|0|325|UNDO_LOG|

+------------+-------+------------+------------------+

5 rows in set(0.00 sec)
```

[illegible]5[illegible]TABLE_NAME[illegible]NULL[illegible]

3.[illegible]

[illegible]2-2[illegible]InnoDB[illegible]redo log buffer[illegible]InnoDB[illegible]innodb_log_buffer_size[illegible]8MB[illegible]

```
mysql[illegible]SHOW VARIABLES LIKE'innodb_log_buffer_size'\G;

***************************1.row***************************
```

```
Variable_name:innodb_log_buffer_size

Value:8388608

1 row in set(0.00 sec)
```

[illegible]8MB[illegible]

❑Master Thread[illegible]

❑[illegible]

❑[illegible]1/2[illegible]

4.[illegible]

[illegible]DBA[illegible]InnoDB[illegible]heap[illegible]innodb_buffer_pool[illegible]frame buffer[illegible]buffer control block[illegible]LRU[illegible]InnoDB[illegible]

2.4 Checkpoint[illegible]

[illegible]CPU[illegible]DML[illegible]Update[illegible]Delete[illegible]

[illegible]Write Ahead Log[illegible]ACID[illegible]D[illegible]Durability[illegible]

[illegible]

❑[illegible]

❑[illegible]

[illegible]3TB[illegible]MySQL[illegible]3 TB[illegible]Oracle Exadata[illegible]2 TB[illegible]

[illegible]DBA[illegible]SA[illegible]

[illegible]

[illegible]Checkpoint[illegible]

❑[illegible]

❑[illegible]

❑[illegible]

[illegible]Checkpoint[illegible]Checkpoint[illegible]

[illegible]LRU[illegible]Checkpoint[illegible]

[illegible]Checkpoint[illegible]

[illegible]InnoDB[illegible]LSN[illegible]Log Sequence Number[illegible]LSN[illegible]8[illegible]LSN[illegible]LSN[illegible]Checkpoint[illegible]LSN[illegible]SHOW ENGINE INNODB STATUS[illegible]

```
mysql[illegible]SHOW ENGINE INNODB STATUS\G;

......

---

LOG

---

Log sequence number 92561351052

Log flushed up to 92561351052

Last checkpoint at 92561351052

......
```

[illegible]InnoDB[illegible]Checkpoint[illegible]Checkpoint[illegible]Checkpoint[illegible]InnoDB[illegible]Checkpoint[illegible]

❑Sharp Checkpoint

❑Fuzzy Checkpoint

Sharp Checkpoint□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□innodb_fast_shutdown=1□

□□□□□□□□□□□□□Sharp Checkpoint□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□Fuzzy Checkpoint□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□Fuzzy Checkpoint□

❑Master Thread Checkpoint

❑FLUSH_LRU_LIST Checkpoint

❑Async/Sync Flush Checkpoint

❑Dirty Page too much Checkpoint

□□Master Thread□2.5□□□□□□□□□□□Master Thread□□□□□□□□Checkpoint□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□

FLUSH_LRU_LIST Checkpoint□□□InnoDB□□□□□□□□LRU□□□□□□□□□100□□□□□□□□□□InnoDB1.1.x□□□□□□□□□LRU□□100□□□□□□□□□InnoDB□□□□□□LRU□□□□□□□□□□□□□□□□□□□□□□□□□Checkpoint□□□□□□□□LRU□□□□□□□□FLUSH_LRU_LIST Checkpoint□

□□MySQL 5.6□□□□□□InnoDB1.2.x□□□□□□□□□□□□□□□□□□Page Cleaner□□□□□□□□□□□□□□□□innodb_lru_scan_depth□□□LRU□□□□□□□□□□□□□□□1024□□□

```
mysql□SHOW VARIABLES LIKE'innodb_lru_scan_depth'\G;

**************************1.row**************************
```

```
Variable_name:innodb_lru_scan_depth

Value:1024

1 row in set(0.00 sec)
```

Async/Sync Flush Checkpoint[illegible]LSN[illegible]redo_lsn[illegible]LSN[illegible]checkpoint_lsn[illegible]

```
checkpoint_age=redo_lsn-checkpoint_lsn
```

[illegible]

```
async_water_mark=75%*total_redo_log_file_size

sync_water_mark=90%*total_redo_log_file_size
```

[illegible]1GB[illegible]2GB[illegible]async_water_mark=1.5GB[illegible]sync_water_mark=1.8GB[illegible]

❑[illegible]checkpoint_age[illegible]async_water_mark[illegible]

❑[illegible]async_water_mark[illegible]checkpoint_age[illegible]sync_water_mark[illegible]Async Flush[illegible]Flush[illegible]checkpoint_age[illegible]async_water_mark[illegible]

❑checkpoint_age[illegible]sync_water_mark[illegible]LOAD DATA[illegible]BULK INSERT[illegible]Sync Flush[illegible]Flush[illegible]checkpoint_age[illegible]async_water_mark[illegible]

[illegible]Async/Sync Flush Checkpoint[illegible]InnoDB 1.2.x[illegible]Async Flush Checkpoint[illegible]Sync Flush Checkpoint[illegible]

InnoDB 1.2.x□□□□——□□□MySQL 5.6□□□□□□□□□□□□□□□□□□□□Page Cleaner Thread□□□□□□□□□□□□□□

MySQL□□□□□□□□□□□□□□Flush□□□□□□LRU□□□□□Checkpoint□□□□□□□□□□□□□□□□Async/Sync Flush□□□□□□InnoSQL□□□□□□□□□□□□□□SHOW ENGINE INNODB STATUS□□□□□□

```
mysql□SHOW ENGINE INNODB STATUS\G;
***************************1.row***************************
Type:InnoDB
……
LRU len:112902,unzip_LRU len:0
I/O sum[0]:cur[0],unzip sum[0]:cur[0]
Async Flush:0,Sync Flush:0,LRU List Flush:0,Flush List Flush:111736
……
1 row in set(0.01 sec)
```

□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□9□□□□□

□□□□Checkpoint□□□□Dirty Page too much□□□□□□□□□□□□InnoDB□□□□□□□□Checkpoint□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□innodb_max_dirty_pages_pct□□□

```
mysql□SHOW VARIABLES LIKE'innodb_max_dirty_pages_pct'\G;
***************************1.row***************************
Variable_name:innodb_max_dirty_pages_pct
Value:75
1 row in set(0.00 sec)
```

innodb_max_dirty_pages_pct□□75□□□□□□□□□□□□□□□75%□□□□□□Checkpoint□□□□□□□□□□□□□□InnoDB 1.0.x□□□□□□□□□□□□

90[illegible]75[illegible]

2.5 Master Thread[illegible]

[illegible]2.3[illegible]InnoDB[illegible]Master Thread [illegible]

2.5.1 InnoDB 1.0.x[illegible]Master Thread

Master Thread[illegible]loop[illegible]loop[illegible]backgroup loop[illegible]flush loop[illegible]suspend loop[illegible]Master Thread[illegible]loop[illegible]background loop[illegible]flush loop[illegible]suspendloop[illegible]

Loop[illegible]——[illegible]10[illegible]

```
void master_thread(){

loop[illegible]

for(int i=0;i[illegible]10;i++){

do thing once per second

sleep 1 second if necessary

}

do things once per ten seconds

goto loop;

}
```

[illegible]loop[illegible]thread sleep[illegible]10[illegible]delay[illegible]InnoDB[illegible]

[illegible]

❏[illegible]

❑[illegible]

❑[illegible]100[illegible]InnoDB[illegible]

❑[illegible]background loop[illegible]

[illegible]InnoDB[illegible]commit[illegible]

[illegible]Insert Buffer[illegible]InnoDB[illegible]IO[illegible]5[illegible]5[illegible]InnoDB[illegible]IO[illegible]

[illegible]100[illegible]InnoDB[illegible]buf_get_modified_ratio_pct[illegible]innodb_max_dirty_pages_pct[illegible]90[illegible]90%[illegible]InnoDB[illegible]100[illegible]

[illegible]

```
void master_thread(){

goto loop;

loop[illegible]

for(int i=0;i[illegible]10;i++){

thread_sleep(1)//sleep 1 second

do log buffer flush to disk

if(last_one_second_ios[illegible]5)

do merge at most 5 insert buffer

if(buf_get_modified_ratio_pct[illegible]innodb_max_dirty_pages_pct)

do buffer pool flush 100 dirty page

if(no user activity)

goto backgroud loop

}

do things once per ten seconds
background loop:

do something

goto loop:

}
```

[illegible]10[illegible]

❑[illegible]100[illegible]

❑[illegible]5[illegible]

❑[illegible]

❑[illegible]Undo[illegible]

❑[illegible]100[illegible]10[illegible]

[illegible]InnoDB[illegible]10[illegible]IO[illegible]200[illegible]InnoDB[illegible]IO[illegible]100[illegible]InnoDB[illegible]InnoDB[illegible]

[illegible]InnoDB[illegible]full purge[illegible]Undo[illegible]update[illegible]delete[illegible]consistent read[illegible]full purge[illegible]InnoDB[illegible]undo[illegible]InnoDB[illegible]InnoDB[illegible]full purge[illegible]20[illegible]undo[illegible]

[illegible]InnoDB[illegible]buf_get_modified_ratio_pct[illegible]70%[illegible]100[illegible]70%[illegible]10%[illegible]

[illegible]main loop[illegible]

```
void master_thread(){

goto loop;

loop□

for(int i=0;i□10;i++){

thread_sleep(1)//sleep 1 second

do log buffer flush to disk

if(last_one_second_ios□5)

do merge at most 5 insert buffer

if(buf_get_modified_ratio_pct□innodb_max_dirty_pages_pct)

do buffer pool flush 100 dirty page

if(no user activity)

goto backgroud loop

}

if(last_ten_second_ios□200)

do buffer pool flush 100 dirty page

do merge at most 5 insert buffer

do log buffer flush to disk

do full purge

if(buf_get_modified_ratio_pct□70%)

do buffer pool flush 100 dirty page

else

buffer pool flush 10 dirty page

goto loop

background loop:

do something

goto loop:

}
```

□□□□background loop□□□□□□□□□□□□□□□□□□□□□□□□□
□shutdown□□□□□□□□□□□□background loop□□□□□□□□

- ❑□□□□□Undo□□□□□□

- ❑□□20□□□□□□□□□□

- ❑□□□□□□□□□□□

- ❑□□□□100□□□□□□□□□□□□□□□flush loop□□□□□

□flush loop□□□□□□□□□□□□□InnoDB□□□□□□□□suspend__loop□□Master Thread□□□□□□□□□□□□□□□□□enable□□InnoDB□□□□□□□□□□□□InnoDB□□□□□□□□□Master Thread□□□□□□□□□□

□□□Master Thread□□□□□□□□□

```
void master_thread(){

goto loop;

loop□

for(int i=0;i□10;i++){

thread_sleep(1)//sleep 1 second

do log buffer flush to disk

if(last_one_second_ios□5)

do merge at most 5 insert buffer

if(buf_get_modified_ratio_pct□innodb_max_dirty_pages_pct)

do buffer pool flush 100 dirty page

if(no user activity)

goto backgroud loop

}

if(last_ten_second_ios□200)

do buffer pool flush 100 dirty page

do merge at most 5 insert buffer

do log buffer flush to disk

do full purge
if(buf_get_modified_ratio_pct□70%)
do buffer pool flush 100 dirty page
else
buffer pool flush 10 dirty page
goto loop
background loop:
do full purge
do merge 20 insert buffer
if not idle:
goto loop:
else:
goto flush loop
flush loop:
do buffer pool flush 100 dirty page
if(buf_get_modified_ratio_pct□innodb_max_dirty_pages_pct)
goto flush loop
goto suspend loop
suspend loop:
suspend_thread()
waiting event
goto loop;
}
```

2.5.2 InnoDB1.2.x□□□□□Master Thread

□□□□1.0.x□□□□□Master Thread□□□□□□□□□□□□□□□□□InnoDB□□□□□□IO□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□hard coding□□□□□□□□□□□□□□□□□□□□□SSD□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□IO□□□□□□□□□□□□

□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□100□□□□□□□□□20□□□□□□□□□□□□□□□□□□□□□□□□□□□□100□□□□□□□□□□□□20□□□□□□□□□Master Thread□□□“□□□□”□□□□□□□□□□□□□□□□□□1□□□□□□100□□□□□□20□□□□□□□□□□□□□hard coding□Master Thread□□□□□□□100□□□□□□20□□□insert buffer□□□

□□□□□□□Google□□□□Mark Callaghan□□□□□InnoDB□□□□□□□□□□□□□□□□patch□□InnoDB□□□□□□□□□□□□Google□patch□□□□□□□□□□□□□□□□□□InnoDB Plugin□□InnoDB1.0.x□□□□□□□□□□innodb_io_capacity□□□□□□□IO□□□□□□□□□200□□□□□□□□□□□□□□□□innodb_io_capacity□□□□□□□□□□□□□□□

❑□□□□□□□□□□□□□□□□□□□innodb_io_capacity□□5%□

❑□□□□□□□□□□□□□□□□□□□innodb_io_capacity□

□□□□□□SSD□□□□□□□□□□□□□□RAID□□□□□□□□□□□IO□□□□□□□□□innodb_io_capacity□□□□□□□□□□□□□□IO□□□□□□□

□□□□□□□□□innodb_max_dirty_pages_pct□□□□□□□□InnoDB 1.0.x□□□□□□□□□□□90□□□□□□□□□□□90%□□□□□“□□”□□□□InnoDB□□□□□□□□□□□□□flush loop□□□□□□□□□□□□□□innodb_max_dirty_pages_pct□□□□100□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□20□10□□□□□□□□□□□□□□□□□□innodb_max_dirty_pages_pct□□20□10□□□□□□□□□□□□□□□□□□□□□□□Google□□□□□□□□□□□□□□20□□□□□□□□ [1] □□□InnoDB 1.0.x

□□□□□innodb_max_dirty_pages_pct□□□□□□75□□Google□□□80□□□□□□□□□□□□□□□□□□□□□□□□□□□IO□□□□

InnoDB 1.0.x□□□□□□□□□□□innodb_adaptive_flushing□□□innodb_max_dirty_pages_pct□□□□□□□□□□innodb_max_dirty_pages_pct□□□□100□□□□□□innodb_adaptive_flushing□□□□□□InnoDB□□□□□□□□□□□buf_flush_get_desired_flush_rate□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□buf_flush_get_desired_flush_rate□□□□□□□□□□□redo log□□□□□□□□□□□□□□□□□□□□□□□□□□□□□innodb_max_dirty_pages_pct□□□□□□□□□□□□□

□□□□□□□□□□□□□□full purge□□□□□□□□20□Undo□□□InnoDB 1.0.x□□□□□□□□□innodb_purge_batch_size□□□□□□□□□□full purge□□□Undo□□□□□□□□□□□□□20□□□□□□□□□□□□□□□□□□□

```
mysql□SHOW VARIABLES LIKE'innodb_purge_batch_size'\G;

***************************1.row***************************

Variable_name:innodb_purge_batch_size

Value:20

mysql□SET GLOBAL innodb_purge_batch_size=50;

Query OK,0 rows affected(0.00 sec)
```

□□□□□□□□□□□□□□□□InnoDB 1.0.x□□□□□Master Thread□□□□□□□□□□□□□□□□

```
void master_thread(){

goto loop;

loop□

for(int i=0;i□10;i++){

thread_sleep(1)//sleep 1 second

do log buffer flush to disk
if(last_one_second_ios＜5%innodb_io_capacity)

do merge 5%innodb_io_capacity insert buffer

if(buf_get_modified_ratio_pct＞innodb_max_dirty_pages_pct)

do buffer pool flush 100%innodb_io_capacity dirty page

else if enable adaptive flush

do buffer pool flush desired amount dirty page

if(no user activity)

goto backgroud loop

}

if(last_ten_second_ios＜innodb_io_capacity)

do buffer pool flush 100%innodb_io_capacity dirty page

do merge 5%innodb_io_capacity insert buffer

do log buffer flush to disk

do full purge

if(buf_get_modified_ratio_pct＞70%)

do buffer pool flush 100%innodb_io_capacity dirty page

else

dobuffer pool flush 10%innodb_io_capacity dirty page

goto loop

background loop:

do full purge

do merge 100%innodb_io_capacity insert buffer

if not idle:

goto loop:

else:

goto flush loop

flush loop:

do buffer pool flush 100%innodb_io_capacity dirty page

if(buf_get_modified_ratio_pct＞innodb_max_dirty_pages_pct)

go to flush loop
```

```
goto suspend loop
suspend loop:
suspend_thread()
waiting event
goto loop;
}
```

□□□□□□□□InnoDB 1.0.x□□□□□□□□□□□□□□□□□□□□□□□□□Master Thread□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□Master Thread□□□□□□

□InnoDB 1.0.x□□□□□SHOW ENGINE INNODB STATUS□□□□□□Master Thread□□□□□□□□□□□

```
mysql□SHOW ENGINE INNODB STATUS\G;
***************************1.row***************************
Type:InnoDB
Name:
Status:
=====================================
090921 14:24:56 INNODB MONITOR OUTPUT
=====================================
Per second averages calculated from the last 6 seconds
----------
BACKGROUND THREAD
----------
srv_master_thread loops:45 1_second,45 sleeps,4 10_second,6 background,6 flush
srv_master_thread log flush and writes:45 log writes only:69
……
```

[illegible]45[illegible]sleep[illegible]45[illegible]10[illegible]4[illegible]1:10[illegible]background loop[illegible]6[illegible]flush loop[illegible]6[illegible]MySQL[illegible]

```
mysql[illegible]show engine innodb status\G;

***************************1.row***************************

Type:InnoDB

Name:

Status:

=====================================

091009 10:14:34 INNODB MONITOR OUTPUT

=====================================

Per second averages calculated from the last 42 seconds

----------

BACKGROUND THREAD

----------

srv_master_thread loops:2188 1_second,1537 sleeps,218 10_second,2 background,2 flush

srv_master_thread log flush and writes:1777 log writes only:5816

……
```

[illegible]2188[illegible]sleep[illegible]1537[illegible]InnoDB[illegible]1[illegible]1_second[illegible]sleeps[illegible]

[1] [illegible]http://code.google.com/p/google-mysql-tools/wiki/InnodbIoOltpDisk[illegible]

2.5.3 InnoDB 1.2.x[illegible]Master Thread

[illegible]InnoDB 1.2.x[illegible]Master Thread[illegible]Master Thread[illegible]InnoDB 1.2.x[illegible]Master Thread[illegible]

```
if InnoDB is idle

srv_master_do_idle_tasks();

else

srv_master_do_active_tasks();
```

[illegible]srv_master_do_idle_tasks()[illegible]10[illegible]srv_master_do_active_tasks()[illegible]Master Thread[illegible]Page Cleaner Thread[illegible]Master Thread[illegible]

2.6 InnoDB[illegible]

InnoDB[illegible]

❏[illegible]Insert Buffer[illegible]

❏[illegible]Double Write[illegible]

❏[illegible]Adaptive Hash Index[illegible]

❏[illegible]IO[illegible]Async IO[illegible]

❏[illegible]Flush Neighbor Page[illegible]

[illegible]InnoDB[illegible]

2.6.1 [illegible]

1.Insert Buffer

Insert Buffer[illegible]InnoDB[illegible]InnoDB[illegible]Insert Buffer[illegible]Insert Buffer[illegible]

[illegible]InnoDB[illegible]Primary Key[illegible]SQL[illegible]

```
CREATE TABLE t(

a INT AUTO_INCREMENT,

b VARCHAR(30),

PRIMARY KEY(a)

);
```

[illegible]a[illegible]a[illegible]NULL[illegible]AUTO_INCREMENT[illegible]a[illegible]

[illegible] [illegible]UUID[illegible]NULL[illegible]

[illegible]secondary index[illegible]b[illegible]b[illegible]SQL[illegible]

```
CREATE TABLE t(

a INT AUTO_INCREMENT,

b VARCHAR(30),

PRIMARY KEY(a),

key(b)

);
```

[illegible]a[illegible]b[illegible]B+[illegible]

[illegible]“[illegible]”[illegible]

InnoDB[illegible]Insert Buffer[illegible]Insert Buffer[illegible]Insert Buffer[illegible]merge[illegible]

[illegible]Insert Buffer[illegible]

❑□□□□□□□□secondary index□□

❑□□□□□□□unique□□□

□□□□□□□□□□□InnoDB□□□□□□□Insert Buffer□□Insert Buffer□□□□MySQL□□□□□□□□□□□□□□□□□Insert Buffer□□

□□Insert Buffer□□□□□□

□□□□□□□□SHOW ENGINE INNODB STATUS□□□□□□□□□□□

```
mysql□SHOW ENGINE INNODB STATUS\G;

***************************1.row***************************

Type:InnoDB

Name:

Status:

=====================================

100727 22:21:48 INNODB MONITOR OUTPUT

=====================================

Per second averages calculated from the last 44 seconds

……

-------------------------------------

INSERT BUFFER AND ADAPTIVE HASH INDEX

-------------------------------------

Ibuf:size 7545,free list len 3790,seg size 11336,

8075308 inserts,7540969 merged recs,2246304 merges

……

----------------------------

END OF INNODB MONITOR OUTPUT

============================
1 row in set(0.00 sec)
```

seg size□□□□□Insert Buffer□□□□11336×16KB□□□□177MB□free list len□□□□□□□□□□□size□□□□□□□□□□□□□□□□□□□□□2□□□□□□□□□□□□□□□□□□□□□□□□□□Inserts□□□□□□□□□□merged recs□□□□□□□□□□□□□merges□□□□□□□□□□□□□□□□□□□□merges:merged recs□□□1:3□□□□□□□□□□□□□□□□□□□□IO□□□□□□□□□2/3□

□□□□□□□□□□Insert Buffer□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□innodb_buffer_pool□□□□□□□□□□□1/2□□□□□□□□□□InnoDB□□□□□□□□□□insert buffer□□□□□□□

```
/**Buffer pool size per the maximum insert buffer size*/

#define IBUF_POOL_SIZE_PER_MAX_SIZE 2

ibuf-□max_size=buf_pool_get_curr_size()/UNIV_PAGE_SIZE

/IBUF_POOL_SIZE_PER_MAX_SIZE;
```

□□□□□□□□□□□□□□□□□□□Percona□□□□□patch□□□□□□□□□□□□□□□□□□□□□□□□□Percona□□□□□□□□□□□□□□IBUF_POOL_SIZE_PER_MAX_SIZE□□□□□□□□□□□□□□□□□□□IBUF_POOL_SIZE_PER_MAX_SIZE□□3□□□□□□□□1/3□□□□□□□

2.Change Buffer

InnoDB□1.0.x□□□□□□□Change Buffer□□□□□□Insert Buffer□□□□□□□□□□□□InnoDB□□□□□□□DML□□——INSERT□DELETE□UPDATE□□□□□□□□□□□□Insert Buffer□Delete Buffer□Purge buffer□

□□□□□Insert Buffer□□□Change Buffer□□□□□□□□□□□□□□□□□

□□□□□□□UPDATE□□□□□□□□□□□

❑□□□□□□□□□□

❑□□□□□□□□

□□Delete Buffer□□UPDATE□□□□□□□□□□□□□□□□□□□Purge Buffer□□UPDATE□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□innodb_change_buffering□□□□□□□Buffer□□□□□□□□□□□□□inserts□deletes□purges□changes□all□none□inserts□deletes□purges□□□□□□□□□□□□□changes□□□□inserts□deletes□all□□□□□□□none□□□□□□□□□□□□□□all□

□InnoDB 1.2.x□□□□□□□□□□□innodb_change_buffer_max_size□□□Change Buffer□□□□□□□□□□

```
mysql□SHOW VARIABLES LIKE'innodb_change_buffer_max_size'\G;

***************************1.row***************************

Variable_name:innodb_change_buffer_max_size

Value:25

1 row in set(0.00 sec)
```

innodb_change_buffer_max_size□□□□25□□□□□□□1/4□□□□□□□□□□□□□□□□□□□□□□□□□□□50□

□MySQL 5.5□□□□□□□SHOW ENGINE INNODB STATUS□□□□□□□□□□□□□□

```
mysql□SHOW ENGINE INNODB STATUS\G;

***************************1.row***************************

Type:InnoDB

……

-------------------------------------

INSERT BUFFER AND ADAPTIVE HASH INDEX

-------------------------------------

Ibuf:size 1,free list len 34397,seg size 34399,10875 merges

merged operations:

insert 20462,delete mark 20158,delete 4215
```

```
discarded operations:

insert 0,delete mark 0,delete 0

……
```

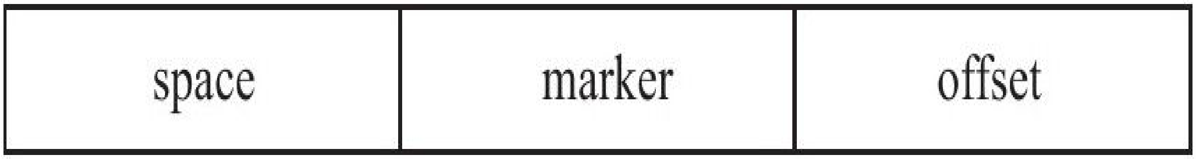
□□□□□□□□□merged operations□discarded operation□□□□□□□□□Change Buffer□□□□□□□□□insert□□Insert Buffer□delete mark□□Delete Buffer□delete□□Purge Buffer□discarded operations□□□Change Buffer□□merge□□□□□□□□□□□□□□□□□□□□□merge□□□□□□□□□

3.Insert Buffer□□□□□

□□□□□□□□□□□□□□□□Insert Buffer□□□□□□□□□□□□□□□□□□□□□□□□Insert Buffer□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□Insert Buffer□□□□□□□□B+□□□MySQL 4.1□□□□□□□□□□□□Insert Buffer B+□□□□□□□□□□□□□□□□□Insert Buffer B+□□□□□□□□□□□□□□□□Insert Buffer□□□□B+□□□□□□□□□□□□□□□□ibdata1□□□□□□□□□□□□□□ibd□□□□□□□□□□□□□□□CHECK TABLE□□□□□□□□□□□□□□□□□□□□□Insert Buffer□□□□□□□□□□□□□□□□ibd□□□□□□□□□□□□□REPAIR TABLE□□□□□□□□□□□□□□□

Insert Buffer□□□B+□□□□□□□□□□□□□□□□□□□□□□□□□□□□□search key□□□□□□□□□□2-3□□□

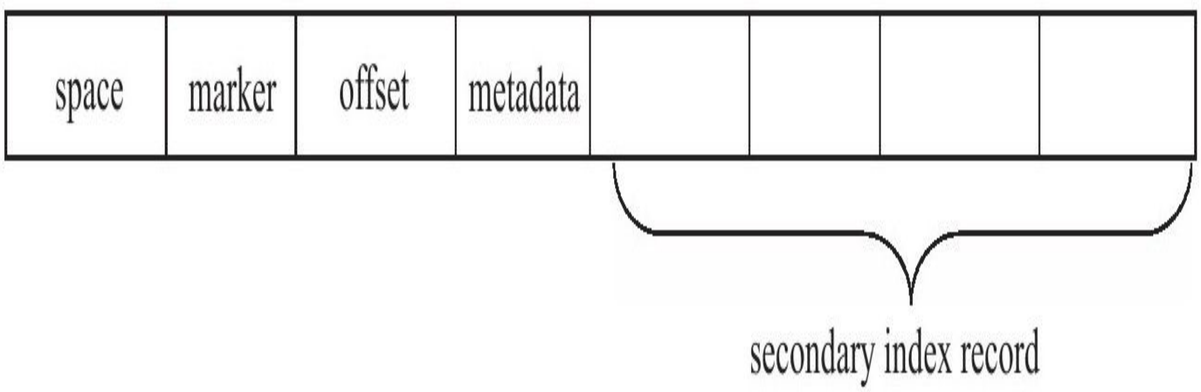


□　2-3　Insert Buffer□□□□□□search key

search key□□□□9□□□□□□space□□□□□□□□□□□□□□id□□InnoDB□□□□□□□□□□□□□□□space id□□□□□space id□□□□□□□□□space□□4□□□marker□□1□□□□□□□□□□□□□Insert Buffer□offset□□□□□□□□□□□□4□□□

□□□□□□□□□□□□□space□offset□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□search key□□□□□□Insert Buffer□□B+□□□□□□□□□□□□□Insert Buffer B+□□□□□□□□

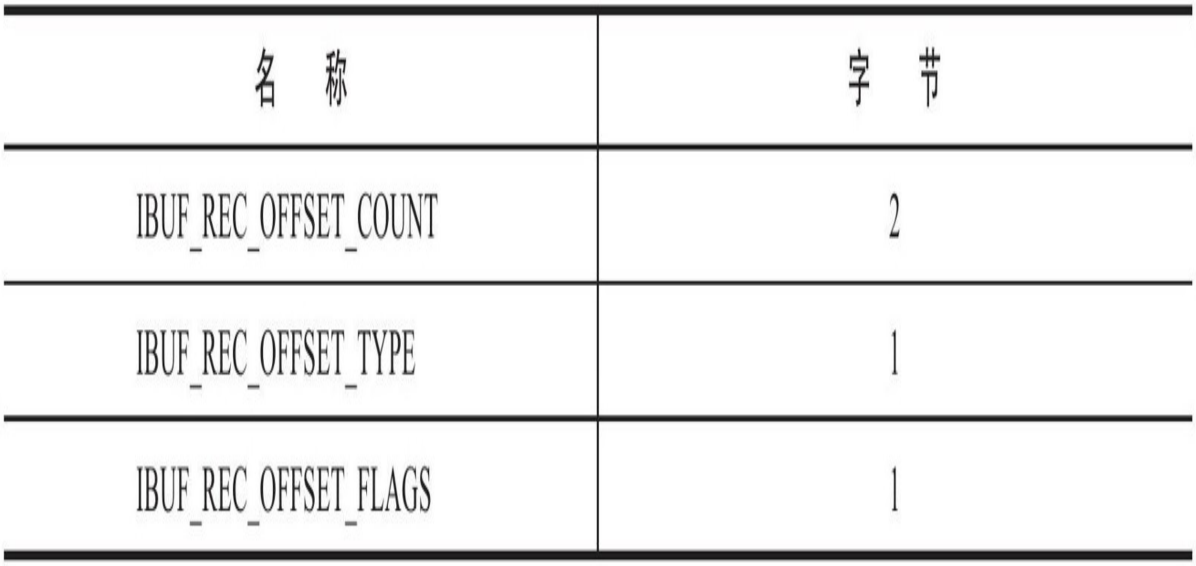
□□□□□Insert Buffer B+□□□□□□□□□□□2-4□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□



□ 2-4 Insert Buffer□□□□□□□□

space□marker□page_no□□□□□□□□□□□□□□□□□□□□9□□□□4□□□metadata□□4□□□□□□□□□□□2-2□□□

表2-2 metadata字段存储的内容



名称	字节
IBUF_REC_OFFSET_COUNT	2
IBUF_REC_OFFSET_TYPE	1
IBUF_REC_OFFSET_FLAGS	1

IBUF_REC_OFFSET_COUNT□□□□□□□□□□□□□□□□□□□□□Insert Buffer□□□□□□□InnoDB1.0.x□□□□Change Buffer□□□□□□□□□□□□Insert Buffer□□□□□□□□□□□□□replay□□□□□□□□□□□□

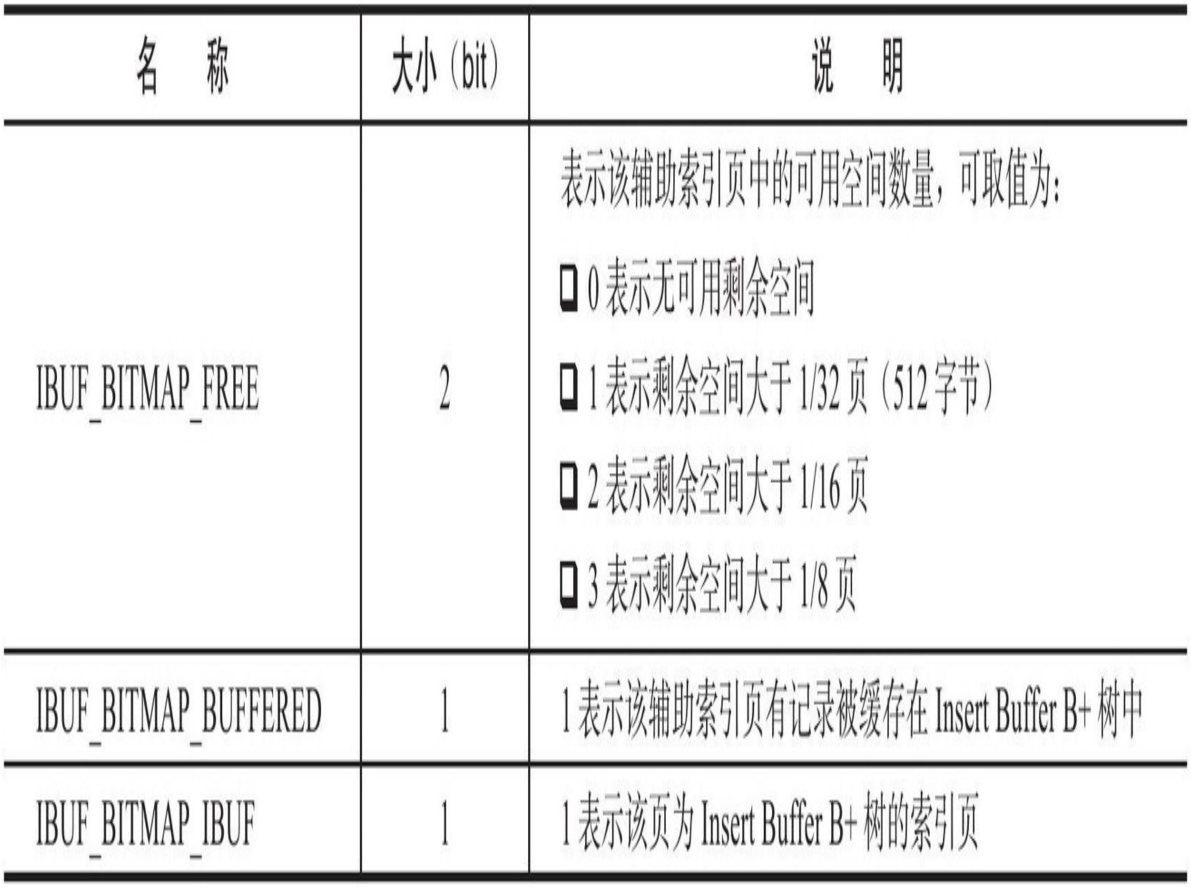
□Insert Buffer□□□□□□5□□□□□□□□□□□□□□□□□□□□□□□□□□□□Insert Buffer B+□□□□□□□□□□□□13□□□□□□

□□□□Insert Buffer□□□□□□□□□□space□page_no□□□□□□□□□□□Insert Buffer B+□□□□□□□□□□□Merge Insert Buffer□□□□□□□□□□□□□□□□□□□□□□□□□□□□space□page_no□□□□□□□□□□□□□□Insert Buffer Bitmap□

□□Insert Buffer Bitmap□□□□□16384□□□□□□□□□□256□□□Extent□□□□Insert Buffer Bitmap□□□16384□□□□□□□□□□□Insert Buffer Bitmap□□□□□□□□□□□□□□□□

□□□□□□□□Insert Buffer Bitmap□□□□4□□bit□□□□2-3□□□□□□□□□

表2-3 每个辅助索引页在Insert Buffer Bitmap中存储的信息



名称	大小（bit）	说明
IBUF_BITMAP_FREE	2	表示该辅助索引页中的可用空间数量，可取值为： ❑ 0表示无可用剩余空间 ❑ 1表示剩余空间大于1/32页（512字节） ❑ 2表示剩余空间大于1/16页 ❑ 3表示剩余空间大于1/8页
IBUF_BITMAP_BUFFERED	1	1表示该辅助索引页有记录被缓存在Insert Buffer B+树中
IBUF_BITMAP_IBUF	1	1表示该页为Insert Buffer B+树的索引页

4.Merge Insert Buffer

[illegible]Insert/Change Buffer[illegible]B+[illegible]B+[illegible]Insert Buffer[illegible]merge[illegible]

[illegible]Merge Insert Buffer[illegible]

❑[illegible]

❑Insert Buffer Bitmap[illegible]

❑Master Thread[illegible]

[illegible]SELECT[illegible]Insert Buffer Bitmap[illegible]Insert Buffer B+[illegible]Insert Buffer B+[illegible]

Insert Buffer Bitmap[illegible]1/32[illegible]1/32[illegible]Insert Buffer B+[illegible]

[illegible]Master Thread[illegible]Master Thread[illegible]10[illegible]Merge Insert Buffer[illegible]merge[illegible]

[illegible]Master Thread[illegible]merge[illegible]srv_innodb_io_capactiy[illegible]InnoDB[illegible]

[illegible]Insert Buffer B+[illegible]space[illegible]offset[illegible]space[illegible]offset[illegible]Insert Buffer[illegible]InnoDB[illegible]Insert Buffer B+[illegible]space[illegible]merge[illegible]merge[illegible]Insert/Change Buffer[illegible]

2.6.2 □□□

□□□Insert Buffer□□InnoDB□□□□□□□□□□□□□□□doublewrite□□□□□□□InnoDB□□□□□□□□□□□□□□

□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□16KB□□□□□□□4KB□□□□□□□□□□□□□□□□□□□□□□□partial page write□□□InnoDB□□□□□□□doublewrite□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□DBA□□□800□□'aaaa'□□□apply□□□doublewrite□□InnoDB□□□□□doublewrite□□□□□□□2-5□□□

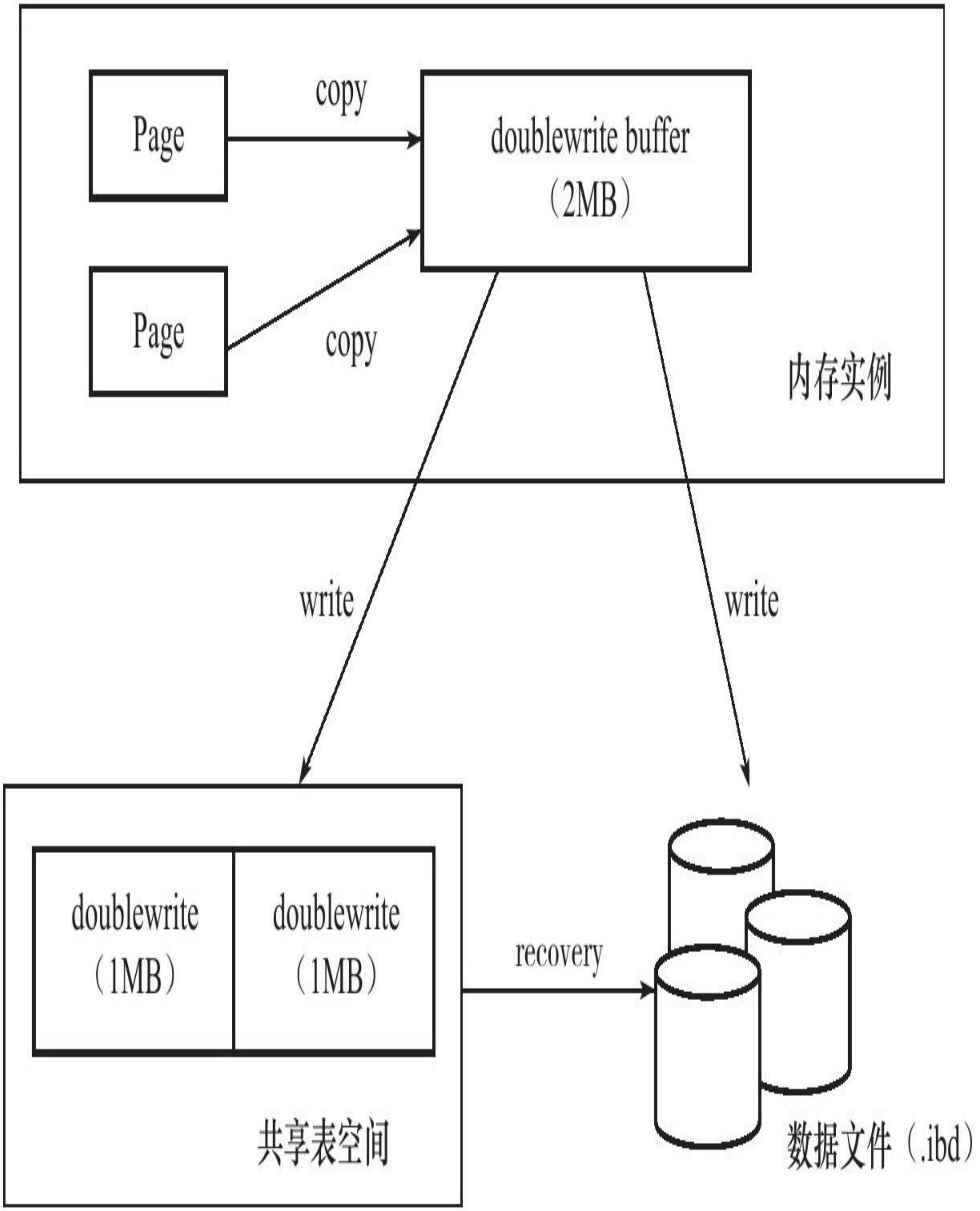


图 2-5 InnoDB存储引擎doublewrite架构

doublewrite□□□□□□□□□□□□□□□doublewrite buffer□□□□2MB□□□□□□□□□□□□□□□□□□□□128□□□□2□□□extent□□□□□□□2MB□□□□□□□□□□□□□□□□□□□□□□□□□□□□memcpy□□□□□□□□□□□□□doublewrite buffer□□□□□doublewrite buffer□□□□□□□1MB□□□□□□□□□□□□□□□□□□□□□□□fsync□□□□□□□□□□□□□□□□□□□□□□□□□□□□□doublewrite□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□doublewrite□□□□□□□□doublewrite buffer□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□doublewrite□□□□□□

```
mysql□SHOW GLOBAL STATUS LIKE'innodb_dblwr%'\G;

***************************1.row***************************

Variable_name:Innodb_dblwr_pages_written

Value:6325194

***************************2.row***************************

Variable_name:Innodb_dblwr_writes

Value:100399

2 rows in set(0.00 sec)
```

□□□□□doublewrite□□□□6 325 194□□□□□□□□□□□□100 399□□□□□□64:1□□□□□□□□□□□□Innodb_dblwr_pages_written:Innodb_dblwr_writes□□□64:1□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□doublewrite□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□doublewrite□□□□□□□□

```
090924 11:36:32 mysqld restarted

090924 11:36:33 InnoDB:Database was not shut down normally!

InnoDB:Starting crash recovery.

InnoDB:Reading tablespace information from the.ibd files...

InnoDB:Crash recovery may have failed for some.ibd files!
```

```
InnoDB:Restoring possible half-written data pages from the doublewrite

InnoDB:buffer...
```

[illegible]MySQL[illegible]SHOW GLOBAL STATUS[illegible]Innodb_buffer_pool_pages_flushed[illegible]doublewrite[illegible]Innodb_dblwr_pages_written[illegible]MySQL 5.5.24[illegible]Innodb_buffer_pool_pages_flushed[illegible]Innodb_dblwr_pages_written[illegible]2[illegible]Bug[illegible]MySQL5.5.24[illegible]Innodb_dblwr_pages_written[illegible]MySQL[illegible]

[illegible]skip_innodb_doublewrite[illegible]doublewrite[illegible]slave server[illegible]slaves erver[illegible]RAID0[illegible]master server[illegible]doublewrite[illegible]

[illegible] [illegible]ZFS[illegible]doublewrite[illegible]

2.6.3 [illegible]

[illegible]hash[illegible]O(1)[illegible]B+[illegible]B+[illegible]B+[illegible]3[illegible]4[illegible]3[illegible]4[illegible]

InnoDB[illegible]Adaptive Hash Index[illegible]AHI[illegible]AHI[illegible]B+[illegible]InnoDB[illegible]

AHI[illegible]a[illegible]b[illegible]

❑WHERE a=xxx

❑WHERE a=xxx and b=xxx

[illegible]InonDB[illegible]AHI[illegible]AHI[illegible]

❑[illegible]100[illegible]

❑[illegible]N[illegible]N=[illegible]*1/16

[illegible]InnoDB[illegible]AHI[illegible]2[illegible]5[illegible]AHI[illegible]self-tuning[illegible]DBA[illegible]

[illegible]SHOW ENGINE INNODB STATUS[illegible]AHI[illegible]

```
mysql[illegible]SHOW ENGINE INNODB STATUS\G;

***************************1.row***************************

Status:

=====================================
090922 11:52:51 INNODB MONITOR OUTPUT

====================================

Per second averages calculated from the last 15 seconds

……

------------------------------------

INSERT BUFFER AND ADAPTIVE HASH INDEX

------------------------------------

Ibuf:size 2249,free list len 3346,seg size 5596,

374650 inserts,51897 merged recs,14300 merges

Hash table size 4980499,node heap has 1246 buffer(s)

1640.60 hash searches/s,3709.46 non-hash searches/s

……
```

□□□□□□AHI□□□□□□□□□AHI□□□□□□□□□□□□□AHI□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□SELECT*FROM table WHERE index_col='xxx'□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□non-hash searches/s□□□□□□hash searches:non-hash searches□□□□□□□□□□□□□□□□□

□□AHI□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□SHOW ENGINE INNODB STATUS□□□□□□innodb_adaptive_hash_index□□□□□□□□□□□□□□□AHI□□□□□□

2.6.4 □□IO

□□□□□□□□□□□□□□□□□□□□□□□□IO□Asynchronous IO□AIO□□□□□□□□□□□□InnoDB□□□□□□□□□

□AIO□□□□Sync IO□□□□□□□IO□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□□IO□□IO□□□□□□□□□□□IO□□□□□□IO□□□□□□□□□□□□IO□□□□□□□□□AIO□

AIO□□□□□□□□□□□IO Merge□□□□□□□□□IO□□□1□IO□□□□□□□IOPS□□□□□□□□□□□□□□□space□page_no□□□

□8□6□□□8□7□□□8□8□

□□□□□□□16KB□□□□□IO□□□□3□IO□□□□AIO□□□□□□□□□□□□□□□□□□□□space□page_no□□□□□□□AIO□□□□□□□IO□□□□□8□6□□□□□□48KB□□□

□□□Linux□□□□□□iostat□□□□□□□□□rrqm/s□wrqm/s□□□□

```
avg-cpu:%user%nice%system%iowait%steal%idle

4.70 0.00 1.60 13.20 0.00 80.50

Device:rrqm/s wrqm/s r/s w/s rMB/s wMB/s avgrq-sz avgqu-sz await svctm%util

sdc 3905.67 172.00 6910.33 466.67 168.81 18.15 51.91 19.17 2.59 0.13 97.73
```

□InnoDB1.1.x□□□AIO□□□□□InnoDB□□□□□□□□□□□□□□□□InnoDB 1.1.x□□□InnoDB Plugin□□□□□□□□□□□□AIO□□□□□□Native AIO□□□□□□□□□□□□□MySQL□□□□libaio□□□□□□□□□□□□□□□□□□

```
/usr/local/mysql/bin/mysqld:error while loading shared libraries:libaio.so.1:cannot open shared object file:No such file or directory
```

□□□□□□□Native AIO□□□□□□□□□□□Windows□□□Linux□□□□□
Native AIO□□□□Mac OSX□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□MySQL□□□□□□□□□□□□□□□□□□□□□□□□

□□innodb_use_native_aio□□□□□□□□Native AIO□□Linux□□□□□□
□□□□ON□

```
mysql□SHOW VARIABLES LIKE'innodb_use_native_aio'\G;

***************************1.row***************************

Variable_name:innodb_use_native_aio

Value:ON

1 row in set(0.00 sec)
```

□□□□□□□□□□□Native AIO□□□□□InnoDB□□□□□□□□□□□□□□□□
Native AIO□□□□□□□□□75%□

□InnoDB□□□□□□read ahead□□□□□□□□□AIO□□□□□□□□□□□□□□□□
□□□□□AIO□□□

2.6.5 [illegible]

InnoDB[illegible]Flush Neighbor Page[illegible]InnoDB[illegible]extent[illegible]AIO[illegible]IO[illegible]IO[illegible]

❑[illegible]

❑[illegible]IOPS[illegible]

[illegible]InnoDB[illegible]1.2.x[illegible]innodb_flush_neighbors[illegible]IOPS[illegible]0[illegible]

2.7 启动、关闭与恢复

InnoDB是MySQL数据库的存储引擎之一，因此InnoDB存储引擎的启动和关闭，更准确的是指在MySQL实例的启动过程中对InnoDB存储引擎的处理过程。

在关闭时，参数innodb_fast_shutdown影响着表的存储引擎为InnoDB的行为。该参数可取值为0、1、2，默认值为1。

❑0表示在MySQL数据库关闭时，InnoDB需要完成所有的full purge和merge insert buffer，并且将所有的脏页刷新回磁盘。这需要一些时间，有时甚至需要几个小时来完成。如果在进行InnoDB升级时，必须将这个参数调为0，然后再关闭数据库。

❑1是参数innodb_fast_shutdown的默认值，表示不需要完成上述的full purge和merge insert buffer操作，但是在缓冲池中的一些数据脏页还是会刷新回磁盘。

❑2表示不完成full purge和merge insert buffer操作，也不将缓冲池中的数据脏页写回磁盘，而是将日志都写入日志文件。这样不会有任何事务的丢失，但是下次MySQL数据库启动时，会进行恢复操作（recovery）。

当正常关闭MySQL数据库时，下次的启动应该会非常“正常”。但是如果没有正常地关闭数据库，如用kill命令关闭数据库，在MySQL数据库运行中重启服务器，或者在关闭数据库时，将参数innodb_fast_shutdown设为了2时，下次MySQL数据库启动时都会对InnoDB存储引擎的表进行恢复操作。

参数innodb_force_recovery影响了整个InnoDB存储引擎恢复的状况。该参数值默认为0，代表当发生需要恢复时，进行所有的恢复操作，当不能进行有效恢复时，如数据页发生了corruption，MySQL数据库可能发生宕机（crash），并把错误写入错误日志中去。

但是，在某些情况下，可能并不需要进行完整的恢复操作，因为用户自己知道怎么进行恢复。比如在对一个表进行alter table操作时发生意外了，数据库重启时会对InnoDB表进行回滚操作，对于一个大表来说这需要很长时间，可能是几个小时。这时用户可以自行进行恢复，如可以把表删除，从备份中重新导入数据到表，可能这些操作的速度要远远快于回滚操作。

参数innodb_force_recovery可以设置为6个非零值：1～6。大的数字表示包含了前面所有小数字表示的影响。具体情况如下：

❑1(SRV_FORCE_IGNORE_CORRUPT)[illegible]corrupt[illegible]

❑2(SRV_FORCE_NO_BACKGROUND)[illegible]Master Thread[illegible]Master Thread[illegible]full purge[illegible]crash[illegible]

❑3(SRV_FORCE_NO_TRX_UNDO)[illegible]

❑4(SRV_FORCE_NO_IBUF_MERGE)[illegible]

❑5(SRV_FORCE_NO_UNDO_LOG_SCAN)[illegible]Undo Log[illegible]InnoDB[illegible]

❑6(SRV_FORCE_NO_LOG_REDO)[illegible]

[illegible]innodb_force_recovery[illegible]0[illegible]select[illegible]create[illegible]drop[illegible]insert[illegible]update[illegible]delete[illegible]DML[illegible]

[illegible]session[illegible]1 000[illegible]InnoDB[illegible]

```
mysql[illegible]START TRANSACTION;

Query OK,0 rows affected(0.00 sec)

mysql[illegible]UPDATE Profile SET password='';

Query OK,9587770 rows affected(7 min 55.73 sec)

Rows matched:9999248 Changed:9587770 Warnings:0
```

START TRANSACTION[illegible]auto commit[illegible]UPDATE[illegible]UNDO[illegible]undo log[illegible]kill[illegible]MySQL[illegible]

```
[root@nineyou0-43[illegible]]#ps-ef|grep mysqld

root 28007 1 0 13:40 pts/1 00:00:00/bin/sh./bin/mysqld_safe--datadir=/usr/local/mysql/data--pid-
file=/usr/local/mysql/data/nineyou0-43.pid

mysql 28045 28007 42 13:40 pts/1 00:04:23/usr/local/mysql/bin/mysqld--basedir=/usr/local/mysql--datadir=/usr/local/mysql/data-
-user=mysql--pid-file=/usr/local/mysql/data/nineyou0-43.pid--skip-external-locking--port=3306--socket=/tmp/mysql.sock
```

```
root 28110 26963 0 13:50 pts/11 00:00:00 grep mysqld

[root@nineyou0-43□]#kill-9 28007

[root@nineyou0-43□]#kill-9 28045
```

□□kill□□□□□□□□□□□□□□□□□MySQL□□□□□□□□□□□UPDATE□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□err□□□□□□□□□□□□□□□□□□□□□

```
090922 13:40:20 InnoDB:Started;log sequence number 6 2530474615

InnoDB:Starting in background the rollback of uncommitted transactions

090922 13:40:20 InnoDB:Rolling back trx with id 0 5281035,8867280 rows to undo

InnoDB:Progress in percents:1090922 13:40:20

090922 13:40:20[Note]/usr/local/mysql/bin/mysqld:ready for connections.

Version:'5.0.45-log'socket:'/tmp/mysql.sock'port:3306 MySQL Community Server(GPL)

2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46
47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88
89 90 91 92 93 94 95 96 97 98 99 100

InnoDB:Rolling back of trx id 0 5281035 completed

090922 13:49:21 InnoDB:Rollback of non-prepared transactions completed
```

□□□□□□□□□□□□□□□innodb_force_recovery□□0□InnoDB□□8867280□□□□□□□□□□□□9□□□

□□□□□□□□□□□□□□□□□□MySQL□□□□□□□□innodb_force_recovery□□3□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
090922 14:26:23 InnoDB:Started;log sequence number 7 2253251193

InnoDB:!!!innodb_force_recovery is set to 3!!!

090922 14:26:23[Note]/usr/local/mysql/bin/mysqld:ready for connections.

Version:'5.0.45-log'socket:'/tmp/mysql.sock'port:3306 MySQL Community Server(GPL)
```

□□□□□“!!!”□InnoDB□□□□□innodb_force_recovery□□□3□□□

2.8 □□

□□□InnoDB□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□InnoDB□□□□□□□□“□□”□□□□□□□□□□□MySQL□□□□□□□□□□InnoDB□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□3□□□□□MySQL□□□□□□MySQL□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□

第3章　文件

本章将分析构成MySQL数据库和InnoDB存储引擎表的各种类型文件。这些文件有以下这些。

❑参数文件：告诉MySQL实例启动时在哪里可以找到数据库文件，并且指定某些初始化参数，这些参数定义了某种内存结构的大小等设置，还会介绍各种参数的类型。

❑日志文件：用来记录MySQL实例对某种条件做出响应时写入的文件，如错误日志文件、二进制日志文件、慢查询日志文件、查询日志文件等。

❑socket文件：当用UNIX域套接字方式进行连接时需要的文件。

❑pid文件：MySQL实例的进程ID文件。

❑MySQL表结构文件：用来存放MySQL表结构定义文件。

❑存储引擎文件：因为MySQL表存储引擎的关系，每个存储引擎都会有自己的文件来保存各种数据。这些存储引擎真正存储了记录和索引等数据。本章主要介绍与InnoDB有关的存储引擎文件。

3.1 [illegible]

[illegible]1[illegible]MySQL[illegible]MySQL[illegible]mysql--help|grep my.cnf[illegible]

MySQL[illegible]Oracle[illegible]Oracle[illegible]mount[illegible]MySQL[illegible]MySQL[illegible]MySQL[illegible]MySQL[illegible]mysql[illegible]

```
090922 16:25:52 mysqld started

090922 16:25:53 InnoDB:Started;log sequence number 8 2801063211

InnoDB:!!!innodb_force_recovery is set to 1!!!

090922 16:25:53[ERROR]Fatal error:Can't open and lock privilege tables:Table'mysql.host'doesn't exist

090922 16:25:53 mysqld ended
```

MySQL[illegible]mysql[illegible]MySQL[illegible]

MySQL[illegible]vi[illegible]emacs[illegible]

3.1.1 [illegible]

[illegible]/[illegible]key/value[illegible]2[illegible]InnoDB[illegible]innodb_buffer_pool_size[illegible]1G[illegible]innodb_buffer_pool_size=1G[illegible]“[illegible]”[illegible]innodb_buffer_pool_size[illegible]“[illegible]”[illegible]1G[illegible]SHOW VARIABLES[illegible]LIKE[illegible]MySQL 5.1[illegible]

[illegible]information_schema[illegible]GLOBAL_VARIABLES[illegible]

```
mysql[illegible]SELECT*FROM

-[illegible]GLOBAL_VARIABLES

-[illegible]WHERE VARIABLE_NAME LIKE'innodb_buffer%'\G;

***************************1.row***************************

VARIABLE_NAME:INNODB_BUFFER_POOL_SIZE

VARIABLE_VALUE:1073741824

1 row in set(0.00 sec)

mysql[illegible]SHOW VARIABLES LIKE'innodb_buffer%'\G;

***************************1.row***************************

Variable_name:innodb_buffer_pool_size

Value:1073741824

1 row in set(0.00 sec)
```

[illegible]GLOBAL_VARIABLES[illegible]SHOW VARIABLES[illegible]MySQL[illegible]

Oracle[illegible]undocumented parameter[illegible]Oracle“[illegible]”[illegible]SQL Server[illegible]DBA[illegible]MySQL[illegible]Oracle[illegible]SQL Server[illegible]

3.1.2 □□□□

MySQL□□□□□□□□□□□□□□

❑□□□dynamic□□□

❑□□□static□□□

□□□□□□□□□□MySQL□□□read only□□□□□□□SET□□□□□□□□□□□□□□SET□□□□□□

```
SET

|[global|session]system_var_name=expr

|[@@global.|@@session.|@@]system_var_name=expr
```

□□□□□□global□session□□□autocommit□□□□□□□□□□□□□□□□□□□□□□□□□□□□binlog_cache_size□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□read_buffer_size□□□□□□

```
mysql□SET read_buffer_size=524288;

Query OK,0 rows affected(0.00 sec)

mysql□SELECT@@session.read_buffer_size\G;

***************************1.row***************************

@@session.read_buffer_size:524288

1 row in set(0.00 sec)

mysql□SELECT@@global.read_buffer_size\G;

***************************1.row***************************

@@global.read_buffer_size:2093056

1 row in set(0.00 sec)
```

□□□□□□□□□□□□□read_buffer_size□2MB□□□□512KB□□□□□□□□□□□read_buffer_size□□□□□2MB□□□□□□□□□□□□□□□□MySQL□□□□□read_buffer_size□□□2MB□□□□512KB□□□□□□set global|session□□□□□□□□□□□□□□□□□□□□SET@@globl|@@session□□□□□□□□□

```
mysql□SET@@global.read_buffer_size=1048576;

Query OK,0 rows affected(0.00 sec)

mysql□SELECT@@session.read_buffer_size\G;

***************************1.row***************************

@@session.read_buffer_size:524288

1 row in set(0.00 sec)

mysql□SELECT@@global.read_buffer_size\G;

***************************1.row***************************

@@global.read_buffer_size:1048576

1 row in set(0.00 sec)
```

□□□read_buffer_size□□□□□□1MB□□□□□□□read_buffer_size□□□□512KB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□MySQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□MySQL□□□MySQL□□□□□□□□□□□□□□□□□MySQL□□□□□Dynamic System Variables□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□

```
mysql□SET GLOBAL datadir='/db/mysql';

ERROR 1238(HY000):Variable'datadir'is a read only variable
```

3.2 [illegible]

[illegible]MySQL[illegible]MySQL[illegible]

❏[illegible]error log[illegible]

❏[illegible]binlog[illegible]

❏[illegible]slow query log[illegible]

❏[illegible]log[illegible]

[illegible]DBA[illegible]MySQL[illegible]

3.2.1 [illegible]

[illegible]MySQL[illegible]MySQL DBA[illegible]SHOW VARIABLES LIKE'log_error'[illegible]

```
mysql[illegible]SHOW VARIABLES LIKE'log_error'\G;

***************************1.row***************************

Variable_name:log_error

Value:/mysql_data_2/stargazer.log

1 row in set(0.00 sec)

mysql[illegible]system hostname

stargazer
```

[illegible]stargazer[illegible]startgazer.err[illegible]MySQL[illegible]

```
[root@nineyou0-43 data]#tail-n 50 nineyou0-43.err
090924 11:31:18 mysqld started
090924 11:31:18 InnoDB:Started;log sequence number 8 2801063331
090924 11:31:19[ERROR]Fatal error:Can't open and lock privilege tables:Table'mysql.host'doesn't exist
090924 11:31:19 mysqld ended
```

[illegible]mysql[illegible]warning[illegible]InnoDB[illegible]redo log[illegible]

```
090924 11:39:44 InnoDB:ERROR:the age of the last checkpoint is 9433712,
InnoDB:which exceeds the log group capacity 9433498.
InnoDB:If you are using big BLOB or TEXT rows,you must set the
InnoDB:combined size of log files at least 10 times bigger than the
InnoDB:largest such row.
090924 11:40:00 InnoDB:ERROR:the age of the last checkpoint is 9433823,
InnoDB:which exceeds the log group capacity 9433498.
InnoDB:If you are using big BLOB or TEXT rows,you must set the
InnoDB:combined size of log files at least 10 times bigger than the
InnoDB:largest such row.
090924 11:40:16 InnoDB:ERROR:the age of the last checkpoint is 9433645,
InnoDB:which exceeds the log group capacity 9433498.
InnoDB:If you are using big BLOB or TEXT rows,you must set the
InnoDB:combined size of log files at least 10 times bigger than the
InnoDB:largest such row.
```

3.2.2 □□□□□

3.2.1□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□slow log□□□□DBA□□□□□□□□□SQL□□□□□□□SQL□□□□□□□□□□□□□□MySQL□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□□DBA□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□long_query_time□□□□□□□□10□□□10□□

□□□□□□□MySQL□□□□□□□□□□□□□□□□□□□□□□□□□□ON□

```
mysql□SHOW VARIABLES LIKE'long_query_time'\G;

***************************1.row***************************

Variable_name:long_query_time

Value:10.000000

1 row in set(0.00 sec)

mysql□SHOW VARIABLES LIKE'log_slow_queries'\G;

***************************1.row***************************

Variable_name:log_slow_queries

Value:ON

1 row in set(0.00 sec)
```

□□□□□□□□□□□□□□□long_query_time□□□□□□MySQL□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□□□long_query_time□□□□□□□□□□□□□□□□□□□□□□□□□□□long_query_time□□□□□□□□□□□□MySQL 5.1□□□long_query_time□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□DBA□□□□DBA□□□□□SQL□□□□0.5□□0.05□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□log_queries_not_using_indexes□□□□□□SQL□□□□□□□□□□MySQL□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□log_queries_not_using_indexes□

```
mysql□SHOW VARIABLES LIKE'log_queries_not_using_indexes'\G;

***************************1.row***************************

Variable_name:log_queries_not_using_indexes

Value:ON

1 row in set(0.00 sec)
```

MySQL 5.6.5□□□□□□□□□□□log_throttle_queries_not_using_indexes□□□□□□□□□□□□□slow log□□□□□□□□SQL□□□□□□□□□□0□□□□□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□slow log□□□□□slow log□□□□□□□□□□□DBA□□□□□□□□□□□

DBA□□□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□MySQL□□□□□□□□□□□□□□□□□□□□□□□SQL□□MySQL□□□□□□mysqldumpslow□□□□□□□□□□DBA□□□□□□

```
[root@nh122-190 data]#mysqldumpslow nh122-190-slow.log

Reading mysql slow query log from nh122-190-slow.log

Count:11 Time=10.00s(110s)Lock=0.00s(0s)Rows=0.0(0),dbother[dbother]@localhost

insert into test.DbStatus select now(),(N-com_select)/(N-uptime),(N-com_insert)/(N-uptime),(N-com_update)/(N-uptime),(N-
com_delete)/(N-uptime),N-(N/N),N-(N/N),N.N/N,N-N/(N*N),GetCPULoadInfo(N)from test.CheckDbStatus order by check_id desc limit N

Count:653 Time=0.00s(0s)Lock=0.00s(0s)Rows=0.0(0),9YOUgs_SC[9YOUgs_SC]@[192.168.43.7]

select custom_name_one from'low_game_schema'.'role_details'where role_id='S'rse and summarize the MySQL slow query log.Options
are

--verbose verbose

--debug debug

--help write this text to standard output

-v verbose

-d debug

-s ORDER what to sort by(al,at,ar,c,l,r,t),'at'is default

al:average lock time

ar:average rows sent
```

```
at:average query time

c:count

l:lock time

r:rows sent

t:query time

-r reverse the sort order(largest last instead of first)

-t NUM just show the top n queries

-a don't abstract all numbers to N and strings to'S'

-n NUM abstract numbers with at least n digits within names

-g PATTERN grep:only consider stmts that include this string

-h HOSTNAME hostname of db server for*-slow.log filename(can be wildcard),

default is'*',i.e.match all

-i NAME name of server instance(if using mysql.server startup script)

-l don't subtract lock time from total time
```

□□□□□□□□□□□□□□□10□SQL□□□□□□□□□□□□

```
[root@nh119-141 data]#mysqldumpslow-s al-n 10 david.log

Reading mysql slow query log from david.log

Count:5 Time=0.00s(0s)Lock=0.20s(1s)Rows=4.4(22),Audition[Audition]@[192.168.30.108]

SELECT OtherSN,State FROM wait_friend_info WHERE UserSN=N

Count:1 Time=0.00s(0s)Lock=0.00s(0s)Rows=1.0(1),audition-kr[audition-kr]@[192.168.30.105]

SELECT COUNT(N)FROM famverifycode WHERE UserSN=N AND verifycode='S'

……
```

MySQL 5.1□□□mysql□□□□□□slow_log□□□□□□□□□□

```
mysql□SHOW CREATE TABLE mysql.slow_log\G;

***************************1.row***************************
```

```
Table:slow_log

Create Table:CREATE TABLE'slow_log'(

'start_time'timestamp NOT NULL DEFAULT CURRENT_TIMESTAMP ON UPDATE CURRENT_TIMESTAMP,

'user_host'mediumtext NOT NULL,

'query_time'time NOT NULL,

'lock_time'time NOT NULL,

'rows_sent'int(11)NOT NULL,

'rows_examined'int(11)NOT NULL,

'db'varchar(512)NOT NULL,

'last_insert_id'int(11)NOT NULL,

'insert_id'int(11)NOT NULL,

'server_id'int(11)NOT NULL,

'sql_text'mediumtext NOT NULL

)ENGINE=CSV DEFAULT CHARSET=utf8 COMMENT='Slow log'

1 row in set(0.00 sec)
```

□□log_output□□□□□□□□□□□□□□□FILE□□□□□□□TABLE□□□□□□□□mysql□□□□slow_log□□□□□

```
mysql□SHOW VARIABLES LIKE'log_output'\G;

+---------------+---------+

|Variable_name|Value|

+---------------+---------+

|log_output|FILE|

+---------------+---------+

1 row in set(0.00 sec)

mysql□SET GLOBAL log_output='TABLE';

Query OK,0 rows affected(0.00 sec)

mysql□SHOW VARIABLES LIKE'log_output'\G;

+---------------+---------+
|Variable_name|Value|

+---------------+---------+

|log_output|TABLE|

+---------------+---------+

1 row in set(0.00 sec)

mysql[illegible]select sleep(10)\G;

+-----------+

|sleep(10)|

+-----------+

|0|

+-----------+

1 row in set(10.01 sec)

mysql[illegible]SELECT*FROM mysql.slow_log\G;

***************************1.row***************************

start_time:2009-09-25 13:44:29

user_host:david[david]@localhost[]

query_time:00:00:09

lock_time:00:00:00

rows_sent:1

rows_examined:0

db:mysql

last_insert_id:0

insert_id:0

server_id:0

sql_text:select sleep(10)

1 row in set(0.00 sec)
```

[illegible]log_output[illegible]sleep[illegible]10[illegible]SQL[illegible]slow_log[illegible]

[illegible]slow_log[illegible]CSV[illegible]slow_log[illegible]MyISAM[illegible]start_time[illegible]

```
mysql[illegible]ALTER TABLE mysql.slow_log ENGINE=MyISM;

ERROR 1580(HY000):You cannot'ALTER'a log table if logging is enabled

mysql[illegible]SET GLOBAL slow_query_log=off;

Query OK,0 rows affected(0.00 sec)

mysql[illegible]ALTER TABLE mysql.slow_log ENGINE=MyISAM;

Query OK,1 row affected(0.00 sec)

Records:1 Duplicates:0 Warnings:0
```

[illegible]slow_log[illegible]MyISAM[illegible]

MySQL[illegible]slow log[illegible]SQL[illegible]SQL[illegible]0.5[illegible]

InnoSQL[illegible]SQL[illegible]MySQL[illegible]slow log[illegible]logical reads[illegible]physical reads[illegible]IO[illegible]

```
#Time:111227 23:49:16

#User@Host:root[root]@localhost[127.0.0.1]

#Query_time:6.081214 Lock_time:0.046800 Rows_sent:42 Rows_examined:727558 Logical_reads:91584 Physical_reads:19

use tpcc;

SET timestamp=1325000956;

SELECT orderid,customerid,employeeid,orderdate

FROM orders

WHERE orderdate IN

(SELECT MAX(orderdate)
FROM orders

GROUP BY(DATE_FORMAT(orderdate,'%Y%M'))

);
```

□□□□□□□□□□□□□□□□□□□□□□91 584□□□□□□□19□□□□□□□□□□□□□□□□□SQL□□□□□□□□

□□□□□□□□□□□long_query_io□□□□□□□IO□□□SQL□□□□□slow log□□□□□□□100□□□□□□□□□□□□□□100□SQL□□□□□□slow log□□□□□□□□MySQL□□□□□□□□□□□□□□□slow_query_type□□□□□□□slow log□□□□□□□□□

❑0□□□□SQL□□□□□slow log

❑1□□□□□□□□□SQL□□□□□slow log

❑2□□□□□□IO□□□SQL□□□□□slow log

❑3□□□□□□□□□□□IO□□□SQL□□□□□slow log

3.2.3 □□□□

□□□□□□□□□□MySQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□.log□□□□□□□□□□□

```
[root@nineyou0-43 data]#tail nineyou0-43.log

090925 11:00:24 44 Connect zlm@192.168.0.100 on

44 Query SET AUTOCOMMIT=0

44 Query set autocommit=0

44 Quit

090925 11:02:37 45 Connect Access denied for user'root'@'localhost'(using password:NO)

090925 11:03:51 46 Connect Access denied for user'root'@'localhost'(using password:NO)

090925 11:04:38 23 Query rollback
```

□□□□□□□□□□□□□□□□□□□□□□Access denied□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□MySQL 5.1□□□□□□□□□□□□□□□mysql□□□□general_log□□□□□□□□□□□□□□□□□□slow_log□□□□□□□□□□□□

3.2.4 □□□□□

□□□□□□binary log□□□□□MySQL□□□□□□□□□□□□□□□□□□SELECT□SHOW□□

```
mysql□UPDATE t SET a=1 WHERE a=2;

Query OK,0 rows affected(0.00 sec)

Rows matched:0 Changed:0 Warnings:0

mysql□SHOW MASTER STATUS\G;

***************************1.row***************************

File:mysqld.000008

Position:383

Binlog_Do_DB:

Binlog_Ignore_DB:

Executed_Gtid_Set:

1 row in set(0.00 sec)

mysql□SHOW BINLOG EVENTS IN'mysqld.000008'\G;

***************************1.row***************************

Log_name:mysqld.000008

Pos:4

Event_type:Format_desc

Server_id:1

End_log_pos:120

Info:Server ver:5.6.6-m9-log,Binlog ver:4

***************************2.row***************************

Log_name:mysqld.000008

Pos:120

Event_type:Query

Server_id:1
```

```
End_log_pos:199

Info:BEGIN

***************************3.row***************************

Log_name:mysqld.000008

Pos:199

Event_type:Query

Server_id:1

End_log_pos:303

Info:use'test';UPDATE t SET a=1 WHERE a=2

***************************4.row***************************

Log_name:mysqld.000008

Pos:303

Event_type:Query

Server_id:1

End_log_pos:383

Info:COMMIT

4 rows in set(0.00 sec)
```

[illegible]MySQL[illegible]UPDATE[illegible]Changed[illegible]0[illegible]SHOW BINLOG EVENT[illegible]

[illegible]SELECT[illegible]SHOW[illegible]

❑[illegible]recovery[illegible]point-in-time[illegible]

❑[illegible]replication[illegible]MySQL[illegible]slave[illegible]standby[illegible]MySQL[illegible]master[illegible]primary[illegible]

❑[illegible]audit[illegible]

□□□□□□log-bin[=name]□□□□□□□□□□□□□□□name□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□datadir□□□□

```
mysql□show variables like'datadir';

+---------------+---------------------------+

|Variable_name|Value|

+---------------+---------------------------+

|datadir|/usr/local/mysql/data/|

+---------------+---------------------------+

1 row in set(0.00 sec)

mysql□system ls-lh/usr/local/mysql/data/;

total 2.1G

-rw-rw----1 mysql mysql 6.5M Sep 25 15:13 bin_log.000001

-rw-rw----1 mysql mysql 17 Sep 25 00:32 bin_log.index

-rw-rw----1 mysql mysql 300M Sep 25 15:13 ibdata1

-rw-rw----1 mysql mysql 256M Sep 25 15:13 ib_logfile0

-rw-rw----1 mysql mysql 256M Sep 25 15:13 ib_logfile1

drwxr-xr-x 2 mysql mysql 4.0K May 7 10:08 mysql

drwx------2 mysql mysql 4.0K May 7 10:09 test
```

□□□bin_log.00001□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□bin_log.index□□

□□MySQL□□□□□□□□□□□□□□□□□□□□□□□□1%□□□□□□□□□□□□replication□□point-in-time□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□

❑max_binlog_size

❑binlog_cache_size

❑sync_binlog

❑binlog-do-db

❑binlog-ignore-db

❑log-slave-update

❑binlog_format

□□max_binlog_size□□□+1□□□□□.index□□□□MySQL 5.0□□□□□□□1 073 741 824□□□1 G□□□□□□□max_binlog_size□□□□□1.1G□□

□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□uncommitted□□□□□□□□□□□□□□□□□□□□□□□□□□committed□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□binlog_cache_size□□□□□□□□32K□□□□binlog_cache_size□□□□□□session□□□□□□□□□□□□□□□□□□□□□MySQL□□□□□□□□□□binlog_cache_size□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□binlog_cache_size□□MySQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□SHOW GLOBAL STATUS□□□□binlog_cache_use□binlog_cache_disk_use□□□□□□□□□□binlog_cache_size□□□□□□□□Binlog_cache_use□□□□□□□□□□□□□□□□□binlog_cache_disk_use□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
mysql□show variables like'binlog_cache_size';

+-------------------+-------+

|Variable_name|Value|

+-------------------+-------+

|binlog_cache_size|32768|

+-------------------+-------+

1 row in set(0.00 sec)
```

```
mysql□show global status like'binlog_cache%';

+----------------------+-------------+

|Variable_name|Value|

+----------------------+--------------+

|binlog_cache_disk_use|0|

|binlog_cache_use|33553|

+----------------------+--------------+

2 rows in set(0.00 sec)
```

□□□□□□□33 553□□□□□□□□□□0□□□32KB□□□□□□□□□□□MySQL□□□□□□□□□□□□□□□□binlog_cache_size□□□

□□□sync_binlog=[N]□□□□□□□□□□□□□□□□□□□N□□1□□sync_binlog=1□□□sync_binlog□□□□□0□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ON□□□□□□ON□□□□□□□□□□IO□□□□□□□□□□

□□□□□□sync_binlog□□1□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□COMMIT□□□□□□□sync_binlog□1□□MySQL□□□□□□□□□□□COMMIT□□□innodb_support_xa□□1□□□□□□innodb_support_xa□XA□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□

□□binlog-do-db□binlog-ignore-db□□

□□□□□□□□□□□□slave□□□□□□□□□master□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□log-slave-update□□□□□□□master=□slave=□slave□□□□□□□□□□□□□□□

binlog_format[illegible]MySQL 5.1[illegible]SQL[illegible]statement[illegible]Replication[illegible]Oracle[illegible]Standby[illegible]rand[illegible]uuid[illegible]not sync[illegible]InnoDB[illegible]REPEATABLE READ[illegible]READ COMMITTED[illegible]Oracle[illegible]Microsoft SQL Server[illegible]

MySQL 5.1[illegible]binlog_format[illegible]STATEMENT[illegible]ROW[illegible]MIXED[illegible]

[illegible]1[illegible]STATEMENT[illegible]MySQL[illegible]SQL[illegible]

[illegible]2[illegible]ROW[illegible]SQL[illegible]ROW[illegible]Oracle[illegible]Standby[illegible]Statement[illegible]MySQL 5.1[illegible]binlog_format[illegible]ROW[illegible]InnoDB[illegible]READ COMMITTED[illegible]

[illegible]3[illegible]MIXED[illegible]MySQL[illegible]STATEMENT[illegible]ROW[illegible]

1[illegible]NDB[illegible]DML[illegible]ROW[illegible]

2[illegible]UUID()[illegible]USER()[illegible]CURRENT_USER()[illegible]FOUND_ROWS()[illegible]ROW_COUNT()[illegible]

3[illegible]INSERT DELAY[illegible]

4[illegible]UDF[illegible]

5[illegible]temporary table[illegible]

[illegible]binlog_format[illegible]3-1[illegible]

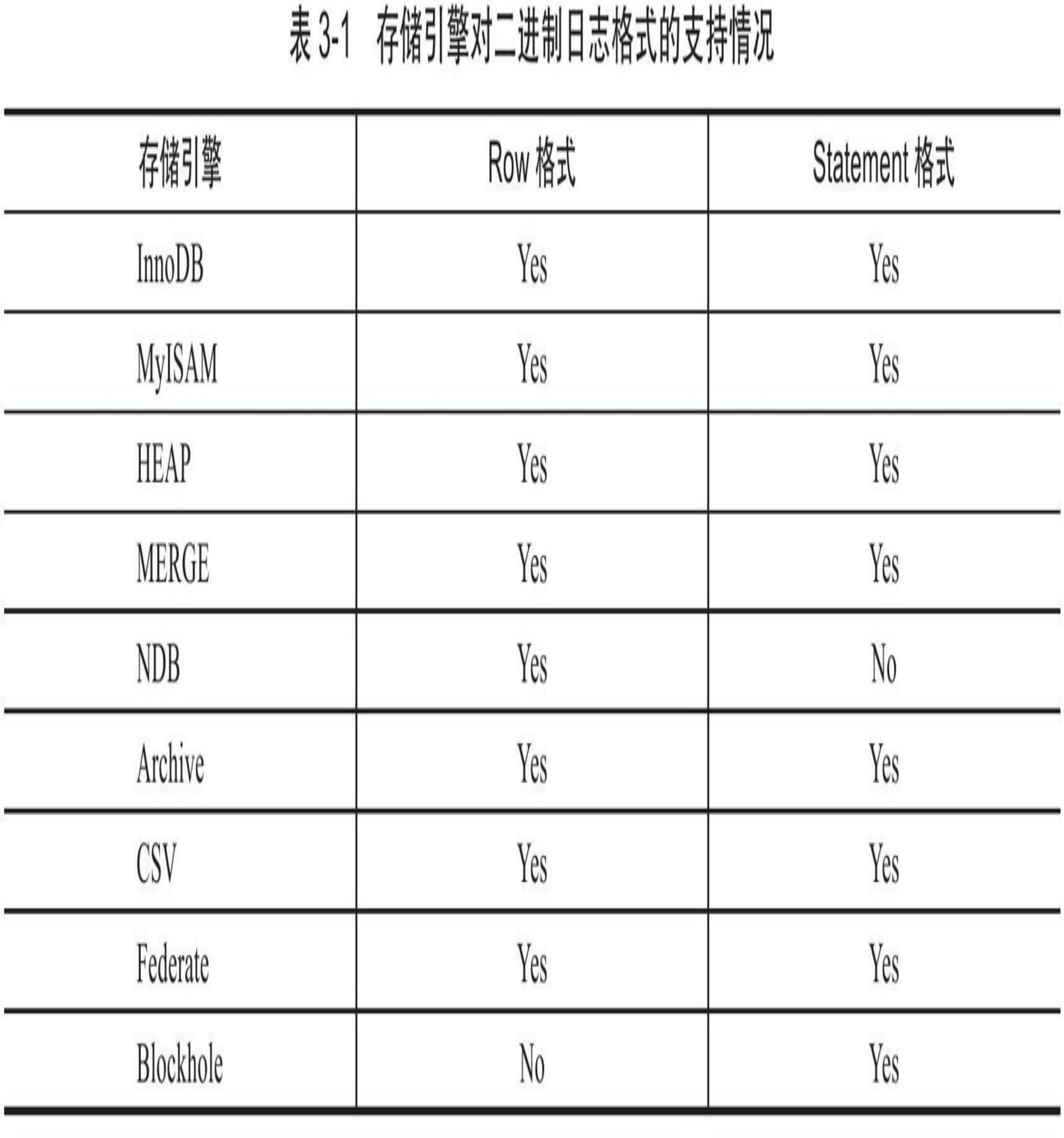


表 3-1 存储引擎对二进制日志格式的支持情况

存储引擎	Row 格式	Statement 格式
InnoDB	Yes	Yes
MyISAM	Yes	Yes
HEAP	Yes	Yes
MERGE	Yes	Yes
NDB	Yes	No
Archive	Yes	Yes
CSV	Yes	Yes
Federate	Yes	Yes
Blockhole	No	Yes

binlog_format□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□binlog_format□□ROW□□□

```
mysql□SET@@session.binlog_format='ROW';

Query OK,0 rows affected(0.00 sec)

mysql□SELECT@@session.binlog_format;
```

```
+------------------------------+

|@@session.binlog_format|

+------------------------------+

|ROW|

+------------------------------+

1 row in set(0.00 sec)
```

□□□□□□□□□□binlog_format□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□

```
mysql□SET GLOBAL binlog_format='ROW';

Query OK,0 rows affected(0.00 sec)

mysql□SELECT@@global.binlog_format;

+-----------------------------+

|@@global.binlog_format|

+-----------------------------+

|ROW|

+-----------------------------+

1 row in set(0.00 sec)
```

□□□□□□□□□□□□binlog_format□□□ROW□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ROW□□□□□□□□□□□□
□□□□□□□□□□□□□□□□100W□□□□□UPDATE□□□□□□□□□□□□□□□□

```
mysql□SELECT@@session.binlog_format\G;

***************************1.row***************************

@@session.binlog_format:STATEMENT

1 row in set(0.00 sec)

mysql□SHOW MASTER STATUS\G;

***************************1.row***************************
```

```
File:test.000003

Position:106

Binlog_Do_DB:

Binlog_Ignore_DB:

1 row in set(0.00 sec)

mysql□UPDATE t1 SET username=UPPER(username);

Query OK,89279 rows affected(1.83 sec)

Rows matched:100000 Changed:89279 Warnings:0

mysql□SHOW MASTER STATUS\G;

***************************1.row***************************

File:test.000003

Position:306

Binlog_Do_DB:

Binlog_Ignore_DB:

1 row in set(0.00 sec)
```

□□□□□□binlog_format□□□STATEMENT□□□□□□□UPDATE□□□□□□□□□□□□□□200□□□306-106□□□□□□ROW□□□□□□t2□□□□□□□□□□□

```
mysql□SET SESSION binlog_format='ROW';

Query OK,0 rows affected(0.00 sec)

mysql□SHOW MASTER STATUS\G;

***************************1.row***************************

File:test.000003

Position:306

Binlog_Do_DB:

Binlog_Ignore_DB:

1 row in set(0.00 sec)

mysql□UPDATE t2 SET username=UPPER(username);

Query OK,89279 rows affected(2.42 sec)
```

```
Rows matched:100000 Changed:89279 Warnings:0

mysql□SHOW MASTER STATUS\G;

***************************1.row***************************

File:test.000003

Position:13782400

Binlog_Do_DB:

Binlog_Ignore_DB:

1 row in set(0.00 sec)
```

□□□□□□□□□□□□□□□ROW□□□□□□□13 782 094□□□□□□□□□□□□□□□□□□□13MB□□□□t2□□□□□□□□17MB□□□□□□□□□□□□□□□□□□□sync_binlog=1□□□□□□□□MySQL□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□binlog_format□□□ROW□□

□□□cat□head□tail□□□□□□□□□□□□□□□□□□□□□□□□□MySQL□□□□□mysqlbinlog□□□STATEMENT□□□□□□□□□□□□□□mysqlbinlog□□□□□□□□□□□□SQL□□□□□

```
[root@nineyou0-43 data]#mysqlbinlog--start-position=203 test.000004

/*!40019 SET@@session.max_insert_delayed_threads=0*/;

….

#090927 15:43:11 server id 1 end_log_pos 376 Query thread_id=188 exec_time=1 error_code=0

SET TIMESTAMP=1254037391/*!*/;

update t2 set username=upper(username)where id=1

/*!*/;

#at 376

#090927 15:43:11 server id 1 end_log_pos 403 Xid=1009

COMMIT/*!*/;
```

```
DELIMITER;
#End of log file
ROLLBACK/*added by mysqlbinlog*/;
/*!50003 SET COMPLETION_TYPE=@OLD_COMPLETION_TYPE*/;
```

[illegible]SQL[illegible]UPDATE t2 SET username=UPPER[illegible]username[illegible] WHERE id=1[illegible]SQL[illegible]mysqlbinlog[illegible]Oracle LogMiner[illegible]ROW[illegible]mysqlbinlog[illegible]“[illegible]”[illegible]unreadable[illegible]

```
[root@nineyou0-43 data]#mysqlbinlog--start-position=1065 test.000004
/*!40019 SET@@session.max_insert_delayed_threads=0*/;
……
#at 1135
#at 1198
#090927 15:53:52 server id 1 end_log_pos 1198 Table_map:'member'.'t2'mapped to number 58
#090927 15:53:52 server id 1 end_log_pos 1378 Update_rows:table id 58 flags:STMT_END_F
BINLOG'
EBq/ShMBAAAAPwAAAK4EAAAAADoAAAAAAAAABm1lbWJlcgACdDIACgMPDw/+CgsPAQwKJAAoAEAA
/gJAAAAA
EBq/ShgBAAAAtAAAAGIFAAAQADoAAAAAAAEACv////8A/AEAAAALYWxleDk5ODh5b3UEOXlvdSA3
Y2JiMzI1MmJhNmI3ZTljNDIyZmFjNTMzNGQyMjA1NAFNLacPAAAAAABjEnpxPBIAAAD8AQAAAAtB
TEVYOTk4OFlPVQQ5eW91IDdjYmIzMjUyYmE2YjdlOWM0MjJmYWM1MzM0ZDIyMDU0AU0tpw8AAAAA
AGMSenE8EgAA
'/*!*/;
#at 1378
#090927 15:53:52 server id 1 end_log_pos 1405 Xid=1110
COMMIT/*!*/;
DELIMITER;
#End of log file
ROLLBACK/*added by mysqlbinlog*/;
```

```
/*!50003 SET COMPLETION_TYPE=@OLD_COMPLETION_TYPE*/;
```

[illegible]SQL[illegible]-v[illegible]-vv[illegible]-vv[illegible]-v[illegible]-vv[illegible]

```
[root@nineyou0-43 data]#mysqlbinlog-vv--start-position=1065 test.000004
……
BINLOG'
EBq/ShMBAAAAPwAAAK4EAAAAADoAAAAAAAAABm1lbWJlcgACdDIACgMPDw/+CgsPAQwKJAAoAEAA
/gJAAAAA
EBq/ShgBAAAAtAAAAGIFAAAQADoAAAAAAAEACv////8A/AEAAAALYWxleDk5ODh5b3UEOXlvdSA3
Y2JiMzI1MmJhNmI3ZTljNDIyZmFjNTMzNGQyMjA1NAFNLacPAAAAAABjEnpxPBIAAAD8AQAAAAtB
TEVYOTk4OFlPVQQ5eW91IDdjYmIzMjUyYmE2YjdlOWM0MjJmYWM1MzM0ZDIyMDU0AU0tpw8AAAAA
AGMSenE8EgAA
'/*!*/;
###UPDATE member.t2
###WHERE
###@1=1/*INT meta=0 nullable=0 is_null=0*/
###@2='david'/*VARSTRING(36)meta=36 nullable=0 is_null=0*/
###@3='family'/*VARSTRING(40)meta=40 nullable=0 is_null=0*/
###@4='7cbb3252ba6b7e9c422fac5334d22054'/*VARSTRING(64)meta=64 nullable=0 is_null=0*/
###@5='M'/*STRING(2)meta=65026 nullable=0 is_null=0*/
###@6='2009:09:13'/*DATE meta=0 nullable=0 is_null=0*/
###@7='00:00:00'/*TIME meta=0 nullable=0 is_null=0*/
###@8=''/*VARSTRING(64)meta=64 nullable=0 is_null=0*/
###@9=0/*TINYINT meta=0 nullable=0 is_null=0*/
###@10=2009-08-11 16:32:35/*DATETIME meta=0 nullable=0 is_null=0*/
###SET
###@1=1/*INT meta=0 nullable=0 is_null=0*/
###@2='DAVID'/*VARSTRING(36)meta=36 nullable=0 is_null=0*/
```

```
###@3=family/*VARSTRING(40)meta=40 nullable=0 is_null=0*/
###@4='7cbb3252ba6b7e9c422fac5334d22054'/*VARSTRING(64)meta=64 nullable=0 is_null=0*/
###@5='M'/*STRING(2)meta=65026 nullable=0 is_null=0*/
###@6='2009:09:13'/*DATE meta=0 nullable=0 is_null=0*/
###@7='00:00:00'/*TIME meta=0 nullable=0 is_null=0*/
###@8=''/*VARSTRING(64)meta=64 nullable=0 is_null=0*/
###@9=0/*TINYINT meta=0 nullable=0 is_null=0*/
###@10=2009-08-11 16:32:35/*DATETIME meta=0 nullable=0 is_null=0*/
#at 1378
#090927 15:53:52 server id 1 end_log_pos 1405 Xid=1110
COMMIT/*!*/;
DELIMITER;
#End of log file
ROLLBACK/*added by mysqlbinlog*/;
/*!50003 SET COMPLETION_TYPE=@OLD_COMPLETION_TYPE*/;
```

[illegible]mysqlbinlog[illegible]update t2 set username=upper(username)where id=1[illegible]10W[illegible]ROW[illegible]13MB[illegible]

3.3 □□□□□

□□□□□□□UNIX□□□□□□□MySQL□□□□UNIX□□□□□□□□□□□□□□□□□□□socket□□□□□□□□□□□□□socket□□□□□□/tmp□□□□□□mysql.sock□

```
mysql□SHOW VARIABLES LIKE'socket'\G;

***************************1.row***************************

Variable_name:socket

Value:/tmp/mysql.sock

1 row in set(0.00 sec)
```

3.4 pid□□

□MySQL□□□□□□□□□□□□□ID□□□□□□□——□□□□□pid□□□□□□□□□□pid_file□□□□□□□□□□□□□□□□□□□□□.pid□

```
mysql□show variables like'pid_file'\G;

***************************1.row***************************

Variable_name:pid_file

Value:/usr/local/mysql/data/xen-server.pid

1 row in set(0.00 sec)
```

3.5 [illegible]

[illegible]MySQL[illegible]MySQL[illegible]MySQL[illegible]frm[illegible]

frm[illegible]v_a[illegible]v_a.frm[illegible]cat[illegible]

```
[root@xen-server test]#cat v_a.frm

TYPE=VIEW

query=select'test'.'a'.'b'AS'b'from'test'.'a'

md5=4eda70387716a4d6c96f3042dd68b742

updatable=1

algorithm=0

definer_user=root

definer_host=localhost

suid=2

with_check_option=0

timestamp=2010-08-04 07:23:36

create-version=1

source=select*from a

client_cs_name=utf8

connection_cl_name=utf8_general_ci

view_body_utf8=select'test'.'a'.'b'AS'b'from'test'.'a'
```

3.6 InnoDB[illegible]

[illegible]MySQL[illegible]InnoDB[illegible]

3.6.1 [illegible]

InnoDB[illegible]tablespace[illegible]10MB[illegible]ibdata1[illegible]tablespace file[illegible]innodb_data_file_path[illegible]

```
innodb_data_file_path=datafle_spec1[;datafle_spec2]...
```

[illegible]

```
[mysqld]

innodb_data_file_path=/db/ibdata1:2000M;/dr2/db/ibdata2:2000M:autoextend
```

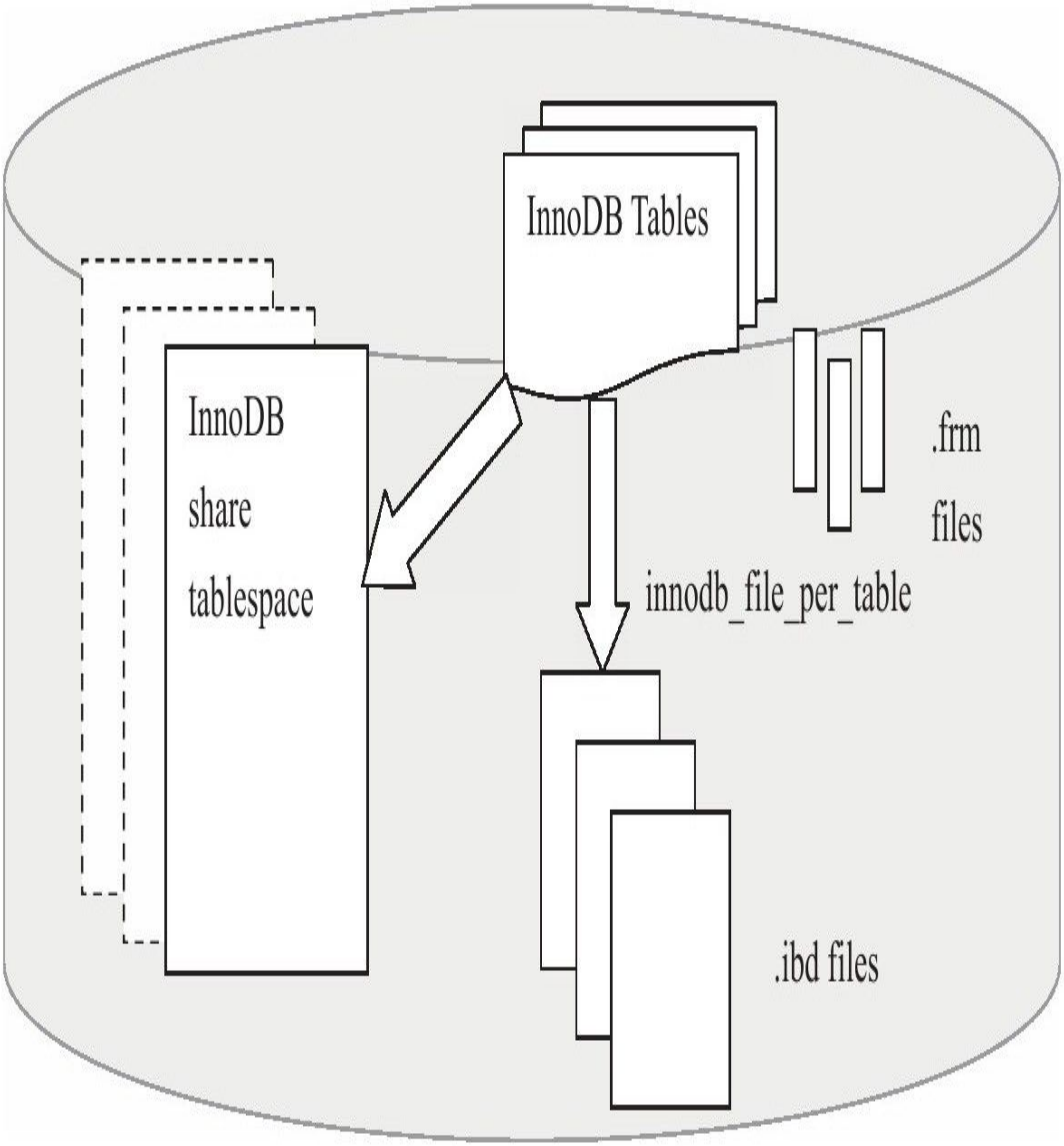
[illegible]/db/ibdata1[illegible]/dr2/db/ibdata2[illegible]idbdata1[illegible]2000MB[illegible]ibdata2[illegible]2000MB[illegible]2000MB[illegible]autoextend[illegible]

[illegible]innodb_data_file_path[illegible]InnoDB[illegible]innodb_file_per_table[illegible]InnoDB[illegible].ibd[illegible]MySQL[illegible]innodb_file_per_table[illegible]

```
mysql[illegible]SHOW VARIABLES LIKE'innodb_file_per_table'\G;

***************************1.row***************************

Variable_name:innodb_file_per_table
```

```
Value:ON

1 row in set(0.00 sec)

mysql□system ls-lh/usr/local/mysql/data/member/*

-rw-r-----1 mysql mysql 8.7K 2009-02-24/usr/local/mysql/data/member/Profile.frm

-rw-r-----1 mysql mysql 1.7G 9□25 11:13/usr/local/mysql/data/member/Profile.ibd

-rw-rw----1 mysql mysql 8.7K 9□27 13:38/usr/local/mysql/data/member/t1.frm

-rw-rw----1 mysql mysql 17M 9□27 13:40/usr/local/mysql/data/member/t1.ibd

-rw-rw----1 mysql mysql 8.7K 9□27 15:42/usr/local/mysql/data/member/t2.frm

-rw-rw----1 mysql mysql 17M 9□27 15:54/usr/local/mysql/data/member/t2.ibd
```

□Profile□t1□t2□□□□InnoDB□□□□□□□□□□□innodb_file_per_table=ON□□□□□□□□□.ibd□□□BITMAP□□□□□□□□□□□□□□□□□□□□□3-1□□□InnoDB□□□□□□□□□□□□□□


InnoDB Tables
InnoDB
share
tablespace
innodb_file_per_table
.frm
files
.ibd files

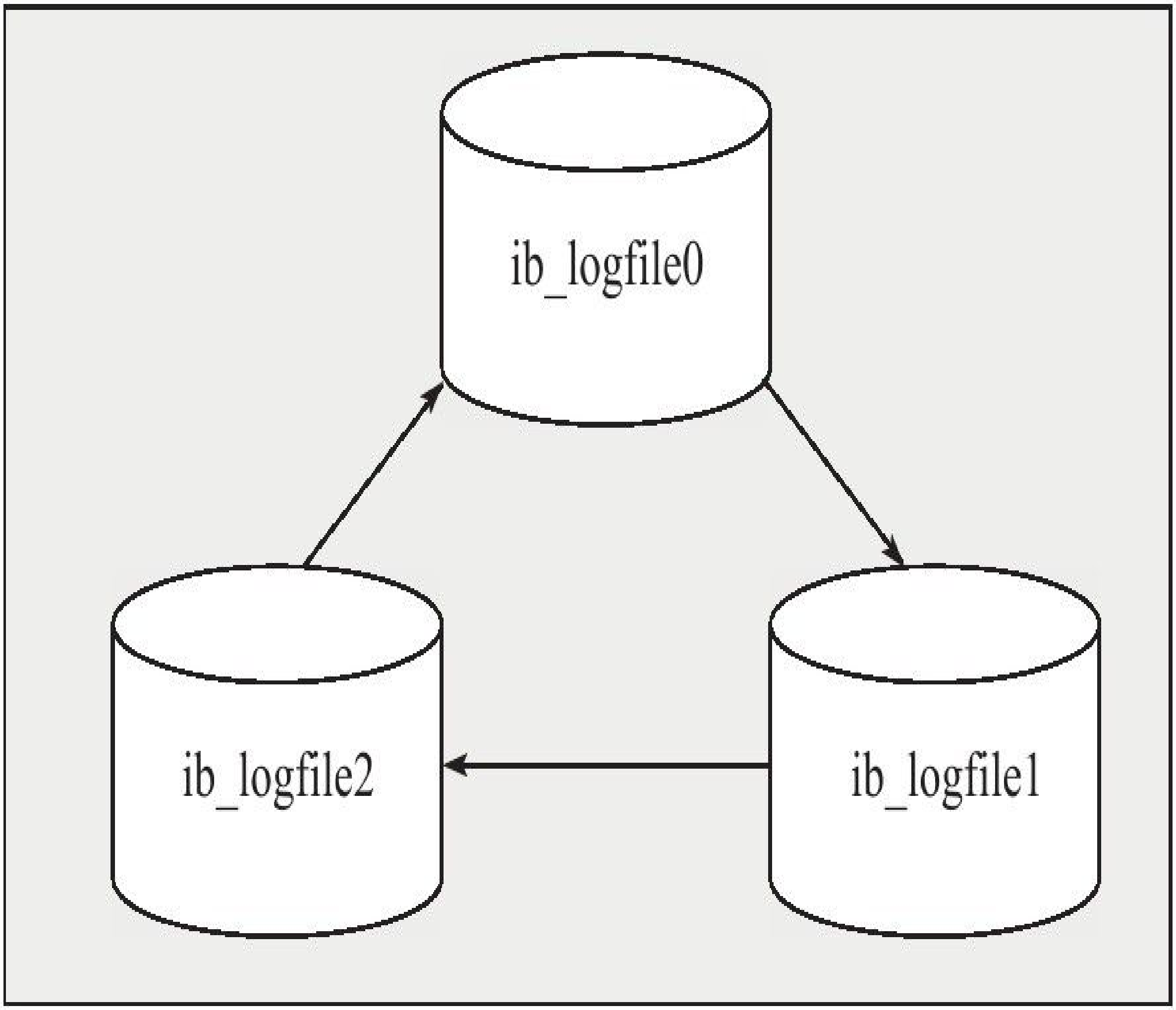
□ 3-1 InnoDB□□□□□□□

3.6.2 □□□□□□

□□□□□□□□InnoDB□□□□□□□□□□□□□□□□ib_logfile0□ib_logfile1□□□□□MySQL□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□redo log file□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□

□□□□□□□□□media failure□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□InnoDB□□□□□□□□1□□□□□□□□group□□□□□□□□□□□2□□□□□□□□□□□□ib_logfile0□ib_logfile1□□□□□□□□□□□□□□□□□□□□□□□□□□□mirrored log groups□□□InnoDB□□□□□□□□□□□□1□□□□□□□□□□□□□□□□□□□□□2□□□□□□□□□2□□□□□□□□□□□□□□□□□1□□□3-2□□□□□□□3□□□□□□□□□□□□□□□□



[illegible] 3-2 [illegible]

[illegible]

❑innodb_log_file_size

❑innodb_log_files_in_group

❑innodb_mirrored_log_groups

❑innodb_log_group_home_dir

□□innodb_log_file_size□□□□□□□□□□□□□□□InnoDB1.2.x□□□□□□□□□□□□□□□□□□□□□4GB□□1.2.x□□□□□□□□□□512GB□

□□innodb_log_files_in_group□□□□□□□□□□□□□□□□□□□□□2□□□innodb_mirrored_log_groups□□□□□□□□□□□□□□□□□1□□innodb_log_group_home_dir□□□□□□□□□□□□□□□□./□□□□MySQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
mysql□SHOW VARIABLES LIKE'innodb%log%'\G;……

***************************4.row***************************

Variable_name:innodb_log_file_size

Value:5242880

***************************5.row***************************

Variable_name:innodb_log_files_in_group

Value:2

***************************6.row***************************

Variable_name:innodb_log_group_home_dir

Value:./

***************************7.row***************************

Variable_name:innodb_mirrored_log_groups

Value:1

7 rows in set(0.00 sec)
```

□□□□□□□□□□□□□InnoDB□□□async checkpoint□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
090924 11:39:44 InnoDB:ERROR:the age of the last checkpoint is 9433712,

InnoDB:which exceeds the log group capacity 9433498.

InnoDB:If you are using big BLOB or TEXT rows,you must set the
```

```
InnoDB:combined size of log files at least 10 times bigger than the
InnoDB:largest such row.
090924 11:40:00 InnoDB:ERROR:the age of the last checkpoint is 9433823,
InnoDB:which exceeds the log group capacity 9433498.
InnoDB:If you are using big BLOB or TEXT rows,you must set the
InnoDB:combined size of log files at least 10 times bigger than the
InnoDB:largest such row.
090924 11:40:16 InnoDB:ERROR:the age of the last checkpoint is 9433645,
InnoDB:which exceeds the log group capacity 9433498.
InnoDB:If you are using big BLOB or TEXT rows,you must set the
InnoDB:combined size of log files at least 10 times bigger than the
InnoDB:largest such row.
```

□□□□□□□InnoDB:ERROR:the age of the last checkpoint is 9433645□InnoDB:which exceeds the log group capacity 9433498□□□□□□□□□□□□capacity□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□innodb buffer pool□□□□□□□flush list□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

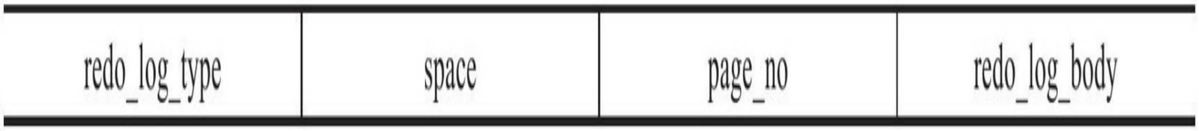
□□□□□□□□□□□□□□MySQL□□□□□□□□□□□□□InnoDB□MyISAM□Heap□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□STATEMENT□□ROW□□□□□MIXED□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□Page□□□□□□□□□□

□□□redo entry□□□□□□□□□□□□□

□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB 1.2.x□□□□□□□□□□51□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□3-2□□□□□□□□□□□□□□

表 3-2 重做日志条目结构



redo_log_type	space	page_no	redo_log_body

□□3-2□□□□□□□□□□□□4□□□□□□

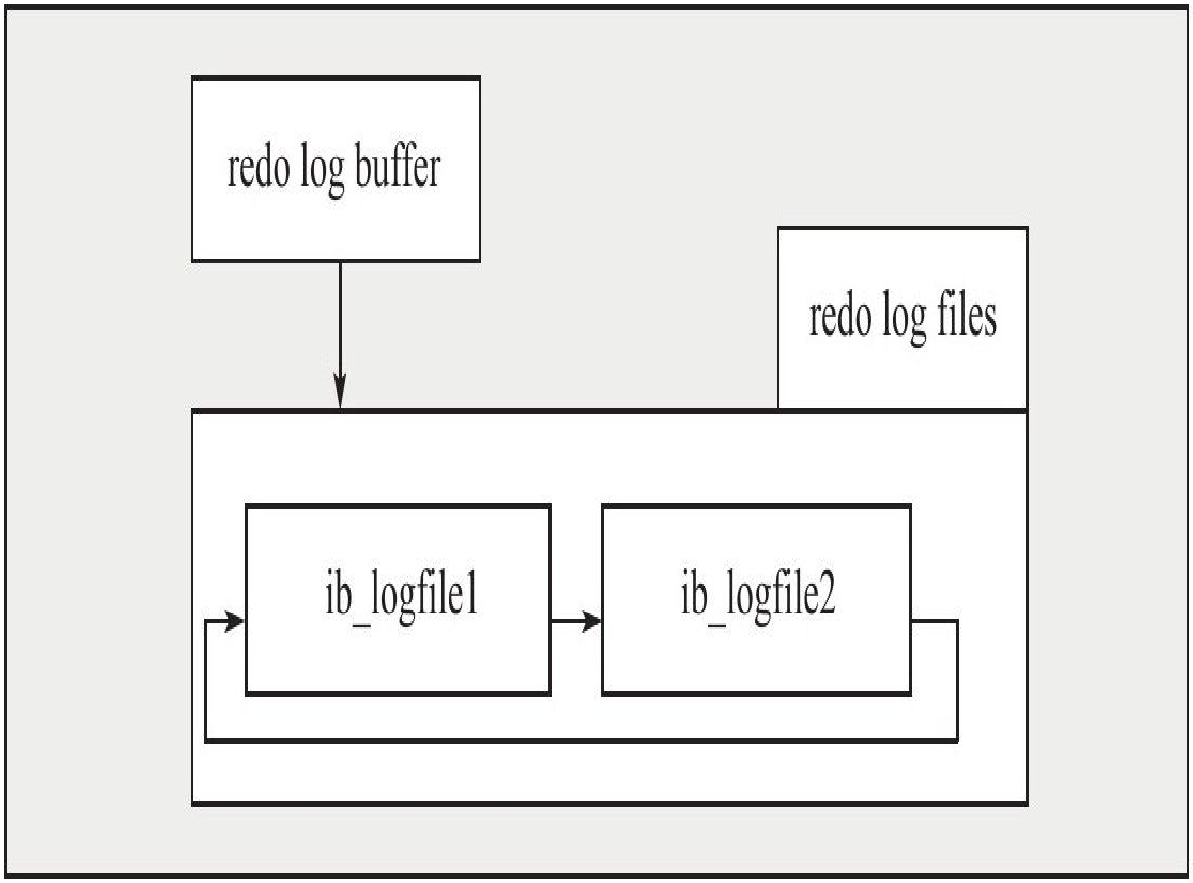
❑redo_log_type□□1□□□□□□□□□□□□

❑space□□□□□□ID□□□□□□□□□□□□□□□□□□□□□4□□

❑page_no□□□□□□□□□□□□□□□□□

❑redo_log_body□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□2□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□redo log buffer□□□□□□□□□□□□□□□□□□□□□□□3-3□□□□□□□□□□□□□□□□



□ 3-3 □□□□□□□

□□□□□□□□□□□□□□□□512□□doublewrite□

□□□2□□□□□□□□master thread□□innodb_flush_log_at_trx_commit□□□□□□□□□commit□□□□□□□□□□□□□□□

□□innodb_flush_log_at_trx_commit□□□□□0□1□2□0□□□1□2□□□□□□□□1□□□□□commit□□□□□□□□□□□□□□□□□□fsync□□□□2□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□commit□□□□□□□□□□□□□□□□□□□□□□□

[illegible] ACID [illegible] innodb_flush_log_at_trx_commit [illegible] 1 [illegible] 0 [illegible] 2 [illegible] 2 [illegible] MySQL [illegible]

3.7 □□

□□□□□□MySQL□□□□□□□□□□□□□□□□□□□□□MySQL□□□□□□□□□□□□□□□□□□□□MySQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□MySQL□□□□□□□□□□□DBA□□DBA□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□point in time□□□□□□□□replication□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□MySQL 5.1□□□□□□□□□□STATEMENT□ROW□MIX□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□DBA□□□□□□□□□□□□□□□□□□□

□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□

[illegible]4[illegible] [illegible]

[illegible]InnoDB[illegible]

4.1 □□□□□

□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□index organized table□□□InnoDB□□□□□□□□□□□□□□□□Primary Key□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□

❑□□□□□□□□□□□□□□□□□Unique NOT NULL□□□□□□□□□□□□□□

❑□□□□□□□□□□InnoDB□□□□□□□□□□6□□□□□□□

□□□□□□□□□□□□□□InnoDB□□□

```
mysql□CREATE TABLE z(

-□a INT NOT NULL,

-□b INT NULL,

-□c INT NOT NULL,

-□d INT NOT NULL,

-□UNIQUE KEY(b),

-□UNIQUE KEY(d),UNIQUE KEY(c));

Query OK,0 rows affected(0.02 sec)

mysql□INSERT INTO z SELECT 1,2,3,4;

Query OK,1 row affected(0.00 sec)

Records:1 Duplicates:0 Warnings:0

mysql□INSERT INTO z SELECT 5,6,7,8;

Query OK,1 row affected(0.00 sec)

Records:1 Duplicates:0 Warnings:0

mysql□INSERT INTO z SELECT 9,10,11,12;

Query OK,1 row affected(0.00 sec)

Records:1 Duplicates:0 Warnings:0
```

□□□□□□□□□□z□□a□b□c□d□□□□b□c□d□□□□□□□□□□□□□□b□□□NULL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□

```
mysql□SELECT a,b,c,d,_rowid FROM z;

+---+------+----+----+--------+

|a|b|c|d|_rowid|

+---+------+----+----+--------+

|1|2|3|4|4|

|5|6|7|8|8|

|9|10|11|12|12|

+---+------+----+----+--------+

3 rows in set(0.00 sec)
```

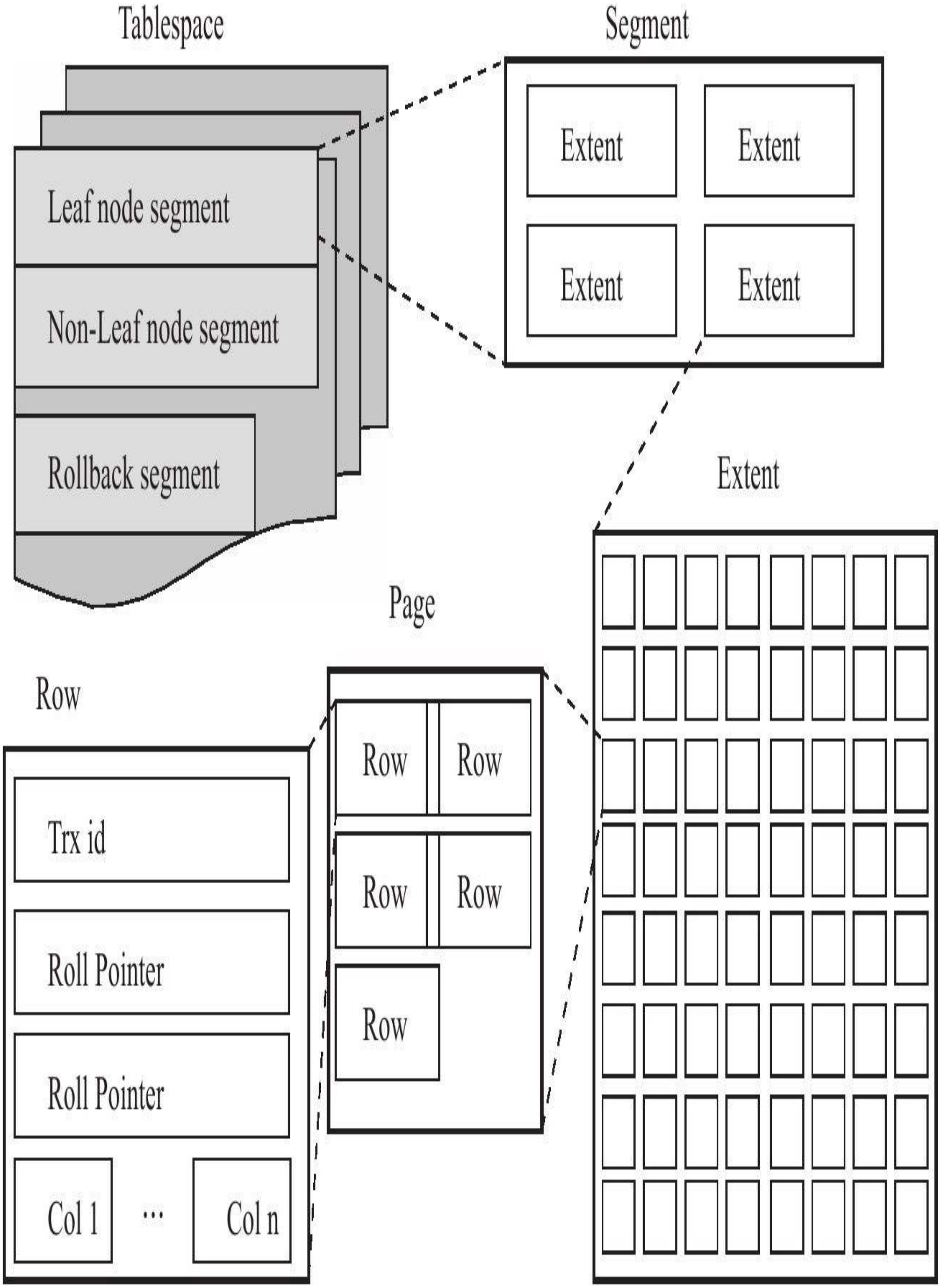
_rowid□□□□□□□□□□□□□□□□□□□□□□z□□□□□□□□□c□d□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□d□□□□□□□□□□□□InnoDB□□□□□□□□□□□

□□□□□□□□□_rowid□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
mysql□CREATE TABLE a(

-□a INT,

-□b INT,

-□PRIMARY KEY(a,b)

-□)ENGINE=InnoDB;

Query OK,0 rows affected(0.03 sec)

mysql□INSERT INTO a SELECT 1,1;

Query OK,1 row affected(0.01 sec)

Records:1 Duplicates:0 Warnings:0

mysql□SELECT a,_rowid FROM a;

ERROR 1054(42S22):Unknown column'_rowid'in'field list'
```

4.2 InnoDB□□□□□□

□InnoDB□□□tablespace□□□□□□□□□segment□□□□extent□□□□page□□□□□□□□□□□□□□□□□□block□□InnoDB□□□□□□□□□□□□□□□4-1□□□



Tablespace
Leaf node segment
Non-Leaf node segment
Rollback segment
Segment
Extent
Extent
Extent
Extent
Extent
Page
Row
Row
Row
Row
Row
Row
Trx id
Roll Pointer
Roll Pointer
Col 1
...
Col n

□ 4-1 InnoDB□□□□□□

4.2.1 □□□

□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□3□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□ibdata1□□□□□□□□□□□□□□□□□□□□□□□□□□innodb_file_per_table□□□□□□□□□□□□□□□□□□□□□□

□□□□□innodb_file_per_table□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Bitmap□□□□□□□□□□□□□undo□□□□□□□□□□□□□□□□□□□□□□□□□Double write buffer□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□innodb_file_per_table□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□innodb_file_per_table□□ON□□□□□□□□□□□□□□□□□□□

```
mysql□SHOW VARIABLES LIKE'innodb_file_per_table'\G;

***************************1.row***************************

Variable_name:innodb_file_per_table

Value:ON

1 row in set(0.00 sec)

mysql□system ls-lh/usr/local/mysql/data/ibdata*

-rw-rw----1 mysql mysql 58M Mar 11 13:58/usr/local/mysql/data/ibdata1
```

□□□□□□□□□□ibdata1□□□□58MB□□□□□□□undo□□□□□□□1□□□□□□mytest□□□□□□□□□□□InnoDB□□□□□□□□

```
mysql□SET autocommit=0;

Query OK,0 rows affected(0.00 sec)

mysql□UPDATE mytest SET salary=0;

Query OK,2844047 rows affected(19.47 sec)

Rows matched:2844047 Changed:2844047 Warnings:0

mysql□system ls-lh/usr/local/mysql/data/ibdata*

-rw-rw----1 mysql mysql 114M Mar 11 14:00/usr/local/mysql/data/ibdata1
```

□□□□□□□□□□□0□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□commit□rollback□□□□□□□□□□□undo□□□□□update mytest set salary=0□□□□□□□□□□□□□□□□ibdata1□□□□□□114MB□□□□□□□□□□□□□□□□□□□□□□□□□□undo□□□

□□□□□□□□□k□□□□□□rollback□ibdata1□□□□□□□□□□□□□□□□□58MB□□□□□□□□□□□□□□□□□□□□□

```
mysql□ROLLBACK;

Query OK,0 rows affected(0.00 sec)

mysql□system ls-lh/usr/local/mysql/data/ibdata*

-rw-rw----1 mysql mysql 114M Mar 11 14:00/usr/local/mysql/data/ibdata1
```

□“□□”□□□□□□□□□□□114MB□□InnoDB□□□□□□□□□rollback□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□undo□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□undo□□□

□□□□□□□2□□□□□master thread□10□□□□□□□full purge□□□□□□□□□□□□□□□□□□□□□□□UPDATE□□□□□□□ibdata1□□□□□□□□□□□□□□□□

□□python□□□□py_innodb_page_info□□□□□□□□□□□□□□□□□□□□□□□□□□code.google.com□□□david-mysql-tools□□□□□□□□□□□□

```
[root@nineyou0-43 py]#python py_innodb_page_info.py/usr/local/mysql/data/ibdata1

Total number of page:83584:

Insert Buffer Free List:204

Freshly Allocated Page:5467

Undo Log Page:38675

File Segment inode:4

B-tree Node:39233

File Space Header:1
```

□□□□□□83 584□□□□□□□□□□□□□□□204□□□5467□□□□□38 675□
undo□□39 233□□□□□□□□□□□□□□-v□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□

4.2.2 □

□4-1□□InnoDB□□□□□□□□□□□□index organized□□□□□□□□□□□□□□□□□□□□□□□B+□□□□□□□□4-1□Leaf node segment□□□□□□□B+□□□□□□□□□4-1□Non-leaf node segment□□□□□□□□□□□□□□□□□□□□□□□□□□

□InnoDB□□□□□□□□□□□□□□□□□□□□□□DBA□□□□□□□□□□□□□□□□Oracle□□□□□□□□□□□□□ASSM□□□□□□□□□□□□□DBA□□□□□□□

4.2.3 □

□□□□□□□□□□□□□□□□□□□□□□□□□□1MB□□□□□□□□□□□□□InnoDB□□□□□□□□□□□4□5□□□□□□□□□□InnoDB□□□□□□□□□16KB□□□□□□□□□64□□□□□□

InnoDB 1.0.x□□□□□□□□□□□□□□□□□□□□□□□KEY_BLOCK_SIZE□□□2K□4K□8K□□□□□□□□□□□□□□□□512□256□128□

InnoDB 1.2.x□□□□□□□innodb_page_size□□□□□□□□□□□□□□□□□□4K□8K□□□□□□□□□□□□□□□□□□□□□□□□□□256□128□□□□□□□□□□□□□□□□□□□□□□1M□

□□□□□□□□□□□□□□□□□□□□□□innodb_file_per_talbe□□□□□□□□□□□96KB□□□□64□□□□□□□□□□□□□□□□1MB□□□□□□□□□□□□□□□□□□□□32□□□□□□□□□fragment page□□□□□□□□□□□□□□□□□□64□□□□□□□□□□□□□□□□□□□□□□□□□□undo□□□InnoDB□□□□□□□□□□□□□

```
mysql□CREATE TABLE t1(

-□col1 INT NOT NULL AUTO_INCREMENT,

-□col2 VARCHAR(7000),

-□PRIMARY KEY(col1))ENGINE=InnoDB;

mysql□system ls-lh/usr/local/mysql/data/test/t1.ibd;

-rw-rw----1 mysql mysql 96K 10□12 14:59/usr/local/mysql/data/test/t1.ibd
```

□□□SQL□□□□□t1□□□col2□□□□VARCHAR□7000□□□□□□□□□□□□□□□2□□□□□□ls□□□□□□□□□□□□□t1□□□□□□□□□□□96KB□□□□□□□SQL□□□

```
mysql□INSERT t1 SELECT NULL,REPEAT('a',7000);

Query OK,1 row affected(0.04 sec)
```

```
Records:1 Duplicates:0 Warnings:0

mysql□INSERT into t1 SELECT NULL,REPEAT('a',7000);

Query OK,1 row affected(0.01 sec)

Records:1 Duplicates:0 Warnings:0

mysql□system ls-lh/usr/local/mysql/data/test/t1.ibd;

-rw-rw----1 mysql mysql 96K 10□12 16:24/usr/local/mysql/data/test/t1.ibd
```

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
py_innodb_page_info□□□□□□□□□□□

```
[root@nineyou0-43 py]#./py_innodb_page_info.py-v/usr/local/mysql/data/test/t1.ibd

page offset 00000000,page type□File Space Header□

page offset 00000001,page type□Insert Buffer Bitmap□

page offset 00000002,page type□File Segment inode□

page offset 00000003,page type□B-tree Node□,page level□0000□

page offset 00000000,page type□Freshly Allocated Page□

page offset 00000000,page type□Freshly Allocated Page□

Total number of page:6:

Freshly Allocated Page:2

Insert Buffer Bitmap:1

File Space Header:1

B-tree Node:1

File Segment inode:1
```

□□□-v□□□□□□□□□□□□□□page offset□3□□□□□□□□□page level□□□□□□□0□□□

```
mysql□INSERT into t1 SELECT NULL,REPEAT('a',7000);

Query OK,1 row affected(0.01 sec)
```

```
Records:1 Duplicates:0 Warnings:0

[root@nineyou0-43 py]#./py_innodb_page_info.py-v/usr/local/mysql/data/test/t1.ibd

page offset 00000000,page type□File Space Header□

page offset 00000001,page type□Insert Buffer Bitmap□

page offset 00000002,page type□File Segment inode□

page offset 00000003,page type□B-tree Node□,page level□0001□

page offset 00000004,page type□B-tree Node□,page level□0000□

page offset 00000005,page type□B-tree Node□,page level□0000□

Total number of page:6:

Insert Buffer Bitmap:1

File Space Header:1

B-tree Node:3

File Segment inode:1
```

□□□□□□page offset□3□□□page level□□□□0□□□1□□□□□□□□□□□□□B+□□□□□□□□□□□□□□□□□B-tree Node□

□□□□□□□□□□□□□□□60□□□□□□□□□□□t1□□□63□□□□32□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
mysql□DELIMITER//

mysql□CREATE PROCEDURE load_t1(count INT UNSIGNED)

-□BEGIN

-□DECLARE s INT UNSIGNED DEFAULT 1;

-□DECLARE c VARCHAR(7000)DEFAULT REPEAT('a',7000);

-□WHILE s□=count DO

-□INSERT INTO t1 SELECT NULL,c;

-□SET s=s+1;

-□END WHILE;

-□END;

-□//
```

```
Query OK,0 rows affected(0.04 sec)

mysql□DELIMITER;

mysql□CALL load_t1(60);

Query OK,1 row affected(1.59 sec)

mysql□SELECT COUNT(*)FROM t1\G;

***************************1.row***************************

count(*):63

1 row in set(0.00 sec)1 row in set(0.00 sec)

mysql□system ls-lh/usr/local/mysql/data/test/t1.ibd;

-rw-rw----1 mysql mysql 576K 10□12 16:56/usr/local/mysql/data/test/t1.ibd
```

□□□□□□□□□63□□□□□□□□□□□□□□□1MB□□□□□□□□□□□□□□□□□□□□□□□□64□□□□□□□□□□□□□py_innodb_page_info□□□□□□□□□t.ibd□□□□□□

```
[root@nineyou0-43 py]#./py_innodb_page_info.py-v/usr/local/mysql/data/test/t1.ibd

page offset 00000000,page type□File Space Header□

page offset 00000001,page type□Insert Buffer Bitmap□

page offset 00000002,page type□File Segment inode□

page offset 00000003,page type□B-tree Node□,page level□0001□

page offset 00000004,page type□B-tree Node□,page level□0000□

page offset 00000005,page type□B-tree Node□,page level□0000□

page offset 00000006,page type□B-tree Node□,page level□0000□

page offset 00000007,page type□B-tree Node□,page level□0000□

page offset 00000008,page type□B-tree Node□,page level□0000□

page offset 00000009,page type□B-tree Node□,page level□0000□

page offset 0000000a,page type□B-tree Node□,page level□0000□

page offset 0000000b,page type□B-tree Node□,page level□0000□

page offset 0000000c,page type□B-tree Node□,page level□0000□

page offset 0000000d,page type□B-tree Node□,page level□0000□
page offset 0000000e,page type□B-tree Node□,page level□0000□

page offset 0000000f,page type□B-tree Node□,page level□0000□

page offset 00000010,page type□B-tree Node□,page level□0000□

page offset 00000011,page type□B-tree Node□,page level□0000□

page offset 00000012,page type□B-tree Node□,page level□0000□

page offset 00000013,page type□B-tree Node□,page level□0000□

page offset 00000014,page type□B-tree Node□,page level□0000□

page offset 00000015,page type□B-tree Node□,page level□0000□

page offset 00000016,page type□B-tree Node□,page level□0000□

page offset 00000017,page type□B-tree Node□,page level□0000□

page offset 00000018,page type□B-tree Node□,page level□0000□

page offset 00000019,page type□B-tree Node□,page level□0000□

page offset 0000001a,page type□B-tree Node□,page level□0000□

page offset 0000001b,page type□B-tree Node□,page level□0000□

page offset 0000001c,page type□B-tree Node□,page level□0000□

page offset 0000001d,page type□B-tree Node□,page level□0000□

page offset 0000001e,page type□B-tree Node□,page level□0000□

page offset 0000001f,page type□B-tree Node□,page level□0000□

page offset 00000020,page type□B-tree Node□,page level□0000□

page offset 00000021,page type□B-tree Node□,page level□0000□

page offset 00000022,page type□B-tree Node□,page level□0000□

page offset 00000023,page type□B-tree Node□,page level□0000□

Total number of page:36:

Insert Buffer Bitmap:1

File Space Header:1

B-tree Node:33

File Segment inode:1
```

□□□□□B-tree Node□□□□33□□□□□□page level□1□□□□□□□□□□32□page level□0□□□□□□□□□□□□□□□□□32□□□□□□□□□□□□□□□□□□□□

[illegible]64[illegible]

```
mysql□CALL load_t1(1);

Query OK,1 row affected(0.10 sec)

mysql□system ls-lh/usr/local/mysql/data/test/t1.ibd;

-rw-rw----1 mysql mysql 2.0M 10□12 17:02/usr/local/mysql/data/test/t1.ibd
```

[illegible]32[illegible] py_innodb_page_info[illegible]t1.ibd[illegible]Freshly Allocated Page[illegible]

```
[root@nineyou0-43 test2]#□/py/py_innodb_page_info.py t1.ibd-v

page offset 00000000,page type□File Space Header□

page offset 00000001,page type□Insert Buffer Bitmap□

page offset 00000002,page type□File Segment inode□

page offset 00000003,page type□B-tree Node□,page level□0001□

page offset 00000004,page type□B-tree Node□,page level□0000□

page offset 00000005,page type□B-tree Node□,page level□0000□

page offset 00000006,page type□B-tree Node□,page level□0000□

page offset 00000007,page type□B-tree Node□,page level□0000□

page offset 00000008,page type□B-tree Node□,page level□0000□

page offset 00000009,page type□B-tree Node□,page level□0000□

page offset 0000000a,page type□B-tree Node□,page level□0000□

page offset 0000000b,page type□B-tree Node□,page level□0000□

page offset 0000000c,page type□B-tree Node□,page level□0000□

page offset 0000000d,page type□B-tree Node□,page level□0000□

page offset 0000000e,page type□B-tree Node□,page level□0000□

page offset 0000000f,page type□B-tree Node□,page level□0000□

page offset 00000010,page type□B-tree Node□,page level□0000□

page offset 00000011,page type□B-tree Node□,page level□0000□
page offset 00000012,page type□B-tree Node□,page level□0000□
page offset 00000013,page type□B-tree Node□,page level□0000□
page offset 00000014,page type□B-tree Node□,page level□0000□
page offset 00000015,page type□B-tree Node□,page level□0000□
page offset 00000016,page type□B-tree Node□,page level□0000□
page offset 00000017,page type□B-tree Node□,page level□0000□
page offset 00000018,page type□B-tree Node□,page level□0000□
page offset 00000019,page type□B-tree Node□,page level□0000□
page offset 0000001a,page type□B-tree Node□,page level□0000□
page offset 0000001b,page type□B-tree Node□,page level□0000□
page offset 0000001c,page type□B-tree Node□,page level□0000□
page offset 0000001d,page type□B-tree Node□,page level□0000□
page offset 0000001e,page type□B-tree Node□,page level□0000□
page offset 0000001f,page type□B-tree Node□,page level□0000□
page offset 00000020,page type□B-tree Node□,page level□0000□
page offset 00000021,page type□B-tree Node□,page level□0000□
page offset 00000022,page type□B-tree Node□,page level□0000□
page offset 00000023,page type□B-tree Node□,page level□0000□
page offset 00000000,page type□Freshly Allocated Page□
……
page offset 00000000,page type□Freshly Allocated Page□
page offset 00000000,page type□Freshly Allocated Page□
page offset 00000000,page type□Freshly Allocated Page□
Total number of page:128:
Freshly Allocated Page:91
Insert Buffer Bitmap:1
File Space Header:1
B-tree Node:34
File Segment inode:1
```

4.2.4 [illegible]

[illegible]InnoDB[illegible]Page[illegible]InnoDB[illegible]InnoDB[illegible]16KB[illegible]InnoDB 1.2.x[illegible]innodb_page_size[illegible]4K[illegible]8K[illegible]16K[illegible]innodb_page_size[illegible]mysqldump[illegible]

[illegible]InnoDB[illegible]

❑[illegible]B-tree Node[illegible]

❑undo[illegible]undo Log Page[illegible]

❑[illegible]System Page[illegible]

❑[illegible]Transaction system Page[illegible]

❑[illegible]Insert Buffer Bitmap[illegible]

❑[illegible]Insert Buffer Free List[illegible]

❑[illegible]Uncompressed BLOB Page[illegible]

❑[illegible]compressed BLOB Page[illegible]

4.2.5 [illegible]

InnoDB[illegible]row-oriented[illegible]16KB/2-200[illegible]7992[illegible]row-oriented[illegible]column-oriented[illegible]MySQL infobright[illegible]SQL[illegible]Sybase IQ[illegible]Google Big Table[illegible]

4.3 InnoDB□□□□□

InnoDB□□□□□□□□□□□□□□□Oracle□Microsoft SQL Server□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB 1.0.x□□□□□InnoDB□□□□□□□Compact□Redundant□□□□□□□□□□□□□□□□□□□□□□□□□□□□Redundant□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□PHYSICAL RECORD□NEW STYLE□□PHYSICAL RECORD□OLD STYLE□□□□□□□□□□□MySQL 5.1□□□□□□□□□Compact□□□□□□□□□□□□SHOW TABLE STATUS LIKE'table_name'□□□□□□□□□□□□□□□row_format□□□□□□□□□□□□□□□□□□□□

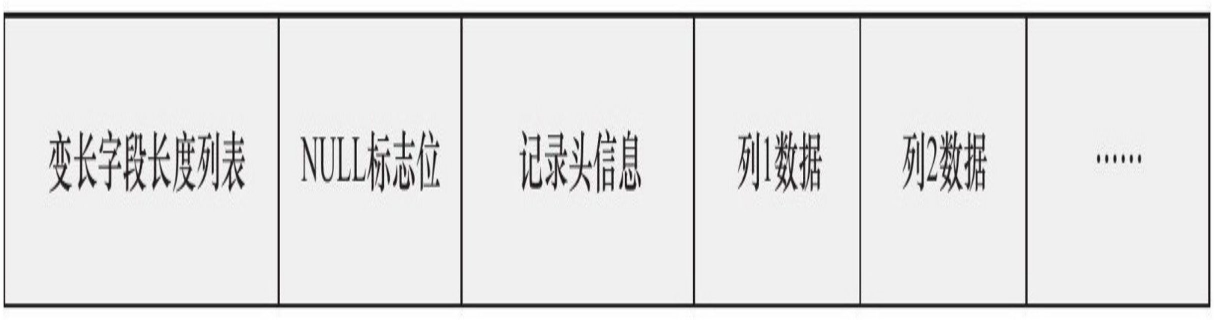
```
mysql□SHOW TABLE STATUS like'mytest%'\G;
***************************1.row***************************
Name:mytest
Engine:InnoDB
Version:10
Row_format:Compact
Rows:6
Avg_row_length:2730
Data_length:16384
Max_data_length:0
Index_length:0
Data_free:0
Auto_increment:NULL
Create_time:2009-03-17 13:33:50
Update_time:NULL
Check_time:NULL
Collation:latin1_swedish_ci
Checksum:NULL
Create_options:
```

```
Comment:

***************************2.row***************************

Name:mytest2

Engine:InnoDB

Version:10

Row_format:Redundant

Rows:0

Avg_row_length:0

Data_length:16384

Max_data_length:0

Index_length:0

Data_free:0

Auto_increment:NULL

Create_time:2009-03-17 13:57:23

Update_time:NULL

Check_time:NULL

Collation:latin1_swedish_ci

Checksum:NULL

Create_options:row_format=REDUNDANT

Comment:

2 rows in set(0.00 sec)
```

[illegible]mytest[illegible]Compact[illegible]mytest2[illegible]Redundant[illegible]py_innodb_page_info[illegible]

4.3.1 Compact[illegible]

Compact[illegible]MySQL 5.0[illegible]4-2[illegible]Compact[illegible]



[illegible] 4-2 Compact[illegible]

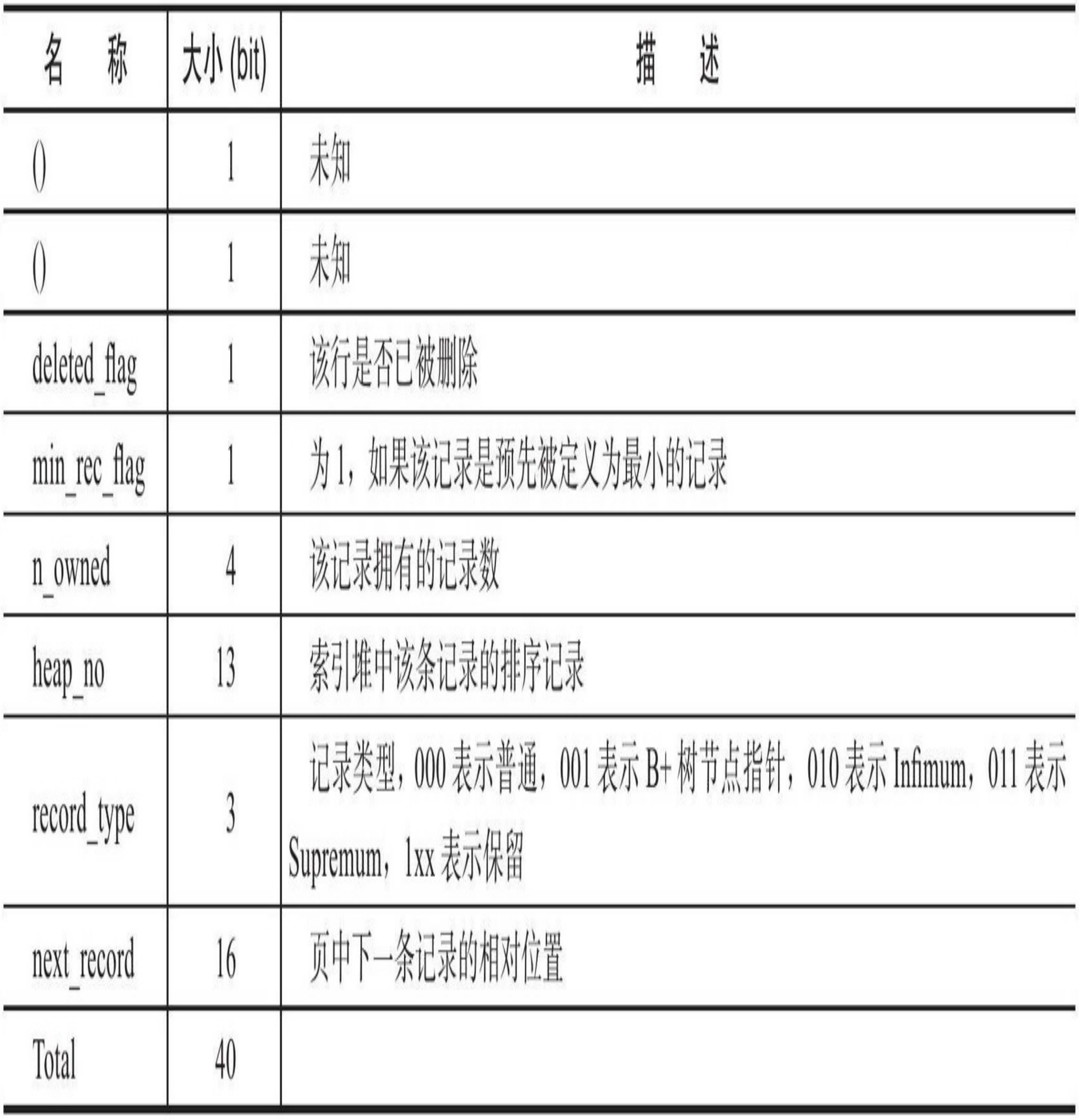
[illegible]4-2[illegible]Compact[illegible]NULL[illegible]

❑[illegible]255[illegible]1[illegible]

❑[illegible]255[illegible]2[illegible]

[illegible]2[illegible]MySQL[illegible]VARCHAR[illegible]65535[illegible]NULL[illegible]NULL[illegible]1[illegible]1[illegible]record header[illegible]5[illegible]40[illegible]4-1[illegible]

表 4-1 Compact 记录头信息



名　称	大小 (bit)	描　述
()	1	未知
()	1	未知
deleted_flag	1	该行是否已被删除
min_rec_flag	1	为 1，如果该记录是预先被定义为最小的记录
n_owned	4	该记录拥有的记录数
heap_no	13	索引堆中该条记录的排序记录
record_type	3	记录类型，000 表示普通，001 表示 B+ 树节点指针，010 表示 Infimum，011 表示 Supremum，1xx 表示保留
next_record	16	页中下一条记录的相对位置
Total	40	

[illegible]NULL[illegible]NULL[illegible]NULL[illegible]ID[illegible]6[illegible]7[illegible]InnoDB[illegible]6[illegible]rowid[illegible]

□□□□□□□□□□□□□Compact□□□□□□□□□

```
mysql□CREATE TABLE mytest(

-□t1 VARCHAR(10),

-□t2 VARCHAR(10),

-□t3 CHAR(10),

-□t4 VARCHAR(10)

-□)ENGINE=INNODB CHARSET=LATIN1 ROW_FORMAT=COMPACT;

Query OK,0 rows affected(0.00 sec)

mysql□INSERT INTO mytest

-□VALUES('a','bb','bb','ccc');

Query OK,1 row affected(0.01 sec)

mysql□INSERT INTO mytest

-□VALUES('d','ee','ee','fff');

Query OK,1 row affected(0.00 sec)

mysql□INSERT INTO mytest

-□VALUES('d',NULL,NULL,'fff');

Query OK,1 row affected(0.00 sec)

mysql□SELECT*FROM mytest\G;

***************************1.row***************************

t1:a

t2:bb

t3:bb

t4:ccc

***************************2.row***************************

t1:d

t2:ee

t3:ee

t4:fff

***************************3.row***************************
```

```
t1:d

t2:NULL

t3:NULL

t4:fff

3 rows in set(0.00 sec)
```

□□□□□□□□□□mytest□□□□□4□□□t1□t2□t4□□VARCHAR□□□□□□□t3□□□□□□□CHAR□□□□□□3□□□□□□□□□□□□□□□□□□□mytest.ibd□□□□□□innodb_file_per_table□□□□□□□□□□□□□□□□□□□□□□ibdata1□□

□Windows□□□□□□□□□□□□□□UltraEdit□□□□□□□□□□Linux□□□□□□□□hexdump-C-v mytest.ibd□mytest.txt□□□□□□□□□□□□□mytest.txt□□□mytest.txt□□□□□□□□□□

```
0000c070 73 75 70 72 65 6d 75 6d 03 02 01 00 00 00 10 00|supremum........|

0000c080 2c 00 00 00 2b 68 00 00 00 00 00 06 05 80 00 00|,...+h..........|

0000c090 00 32 01 10 61 62 62 62 62 20 20 20 20 20 20 20|.2..abbbb|

0000c0a0 20 63 63 63 03 02 01 00 00 00 18 00 2b 00 00 00|ccc........+...|

0000c0b0 2b 68 01 00 00 00 00 06 06 80 00 00 00 32 01 10|+h...........2..|

0000c0c0 64 65 65 65 65 20 20 20 20 20 20 20 20 66 66 66|deeeefff|

0000c0d0 03 01 06 00 00 20 ff 98 00 00 00 2b 68 02 00 00|..........+h...|

0000c0e0 00 00 06 07 80 00 00 00 32 01 10 64 66 66 66 00|........2..dfff.|
```

□□□□□0000c078□□□□□□□□□□□□□□□□□□□□□

```
03 02 01/*□□□□□□□□□□□*/

00/*NULL□□□□□□□□□NULL□*/

00 00 10 00 2c/*Record Header□□□5□□□□*/

00 00 00 2b 68 00/*RowID InnoDB□□□□□6□□*/

00 00 00 00 06 05/*TransactionID*/

80 00 0000 32 01 10/*Roll Pointer*/
```

```
61/*□1□□'a'*/

62 62/*□2□□'bb'*/

62 62 20 20 20 20 20 20 20 20/*□3□□'bb'*/

63 63 63/*□4□□'ccc'*/
```

□□□03 02 01□□□□01 02 03□□□□□□□□InnoDB□□□□□□TransactionID□Roll Pointer□□□□□□□□□□□□CHAR□□□□□□□□□□□□□□□□□□0x20□□□□□□

□□□□□□□Record Header□□□□□□□□□□□□□□□next_recorder□0x2c□□□□□□□□□□□□□□□□□□□□□□□□□0x2c□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□RowID□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□NULL□□□□□□

```
03 01/*□□□□□□□□□□□*/

06/*NULL□□□□□□□□NULL□*/

00 00 20 ff 98/*Record Header*/

00 00 00 2b 68 02/*RowID*/

00 00 00 00 06 07/*TransactionID*/

80 00 00 00 32 01 10/*Roll Pointer*/

64/*□1□□'d'*/

66 66 66/*□4□□'fff'*/
```

□□□□NULL□□□□NULL□□□□□□00□□06□□□□□□□□00000110□□1□□□□□2□□□3□□□□□NULL□□□□□□□□□□□□□□□□□□□□□□NULL□□□□□□□□1□□□4□□NULL□□□□□□□□□□□□□□□□□□□CHAR□□□□VARCHAR□□□□compact□□□NULL□□□□□□□□□□□□

4.3.2　Redundant行记录格式

Redundant是MySQL 5.0版本之前InnoDB的行记录存储方式，MySQL 5.0支持Redundant是为了兼容之前版本的页格式。Redundant行记录采用如图4-3所示的方式存储。

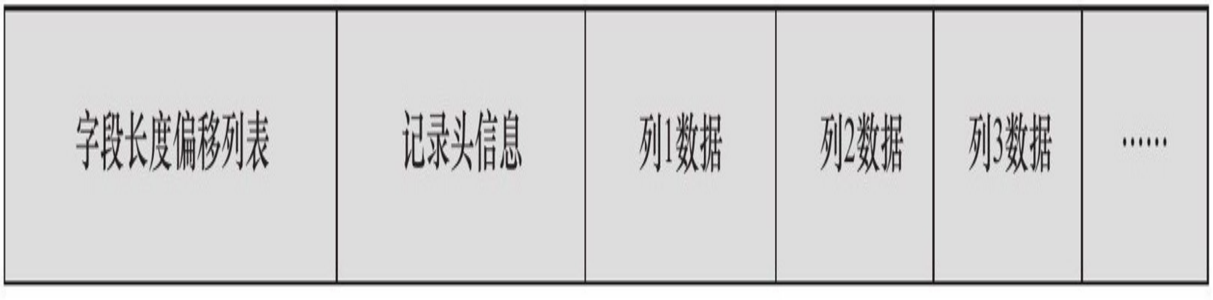


图　4-3　Redundant行记录格式

从图4-3可以看到，不同于Compact行记录格式，Redundant行记录格式的首部是一个字段长度偏移列表，同样是按照列的顺序逆序放置的。若列的长度小于255字节，用1字节来表示；若大于255字节，用2字节来表示。第二个部分为记录头信息（record header），不同于Compact行记录格式，Redundant行记录格式的记录头占用6字节（48位），每位的含义见表4-2。从表4-2中可以发现，n_fields值代表一行中列的数量，占用10位。同时这也很好地解释了为什么MySQL数据库一行支持最多的列为1023。另一个需要注意的值为1byte_offs_flags，该值定义了偏移列表占用1字节还是2字节。而最后的部分就是实际存储每个列的数据。

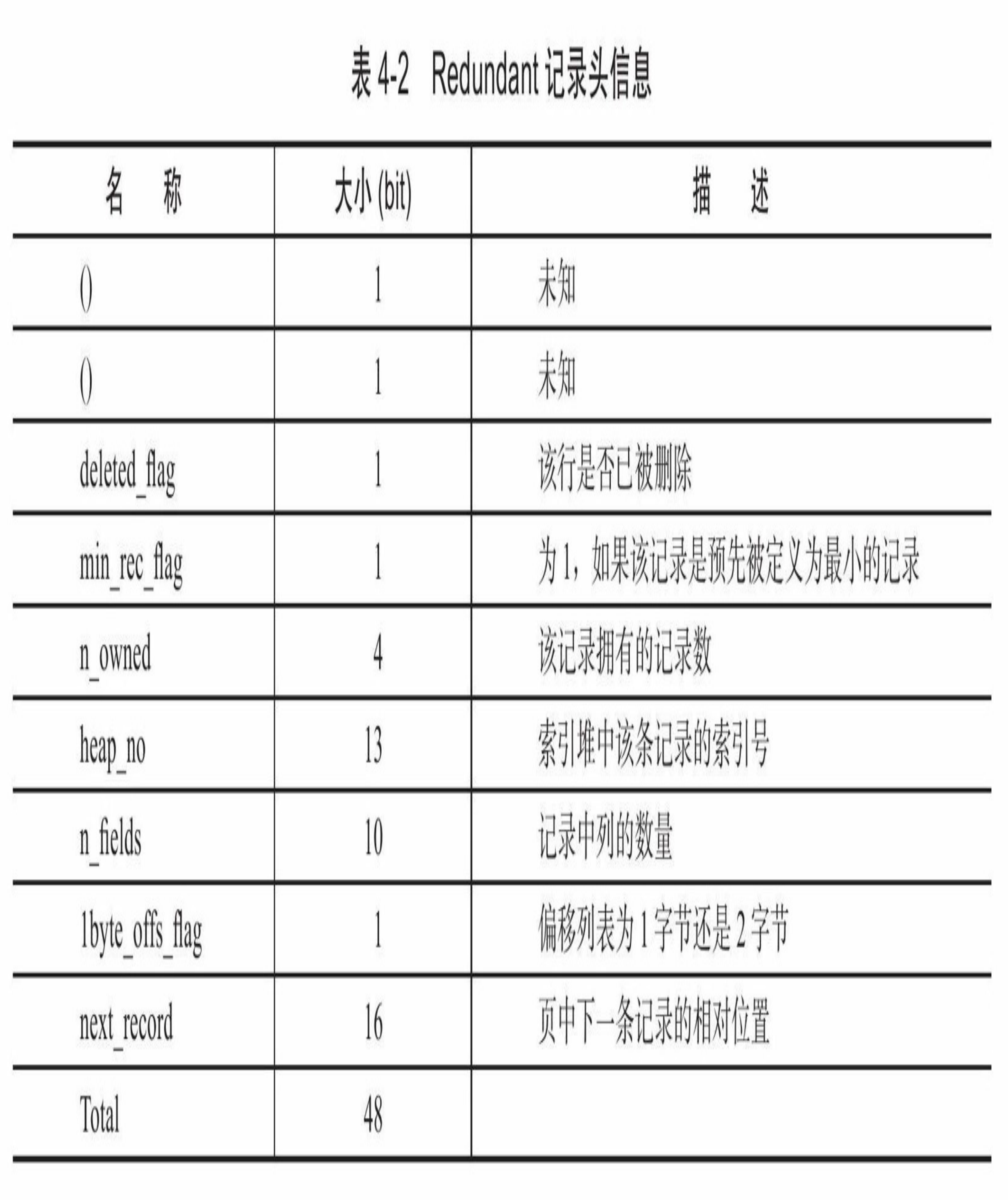


表 4-2 Redundant 记录头信息

名　称	大小 (bit)	描　述
()	1	未知
()	1	未知
deleted_flag	1	该行是否已被删除
min_rec_flag	1	为 1，如果该记录是预先被定义为最小的记录
n_owned	4	该记录拥有的记录数
heap_no	13	索引堆中该条记录的索引号
n_fields	10	记录中列的数量
1byte_offs_flag	1	偏移列表为 1 字节还是 2 字节
next_record	16	页中下一条记录的相对位置
Total	48	

□□□□□□□4.3.1□□□mytest□□□□□□□□□□□Redundant□□mytest2□

```
mysql□CREATE TABLE mytest2

-□ENGINE=InnodB ROW_FORMAT=Redundant
```

```
-□AS

-□SELECT*FROM mytest;

Query OK,3 rows affected(0.00 sec)

Records:3 Duplicates:0 Warnings:0

mysql□SHOW TABLE STATUS LIKE'mytest2'\G;

***************************1.row***************************

Name:mytest2

Engine:InnoDB

Version:10

Row_format:Redundant

Rows:3

Avg_row_length:5461

Data_length:16384

Max_data_length:0

Index_length:0

Data_free:0

Auto_increment:NULL

Create_time:2009-03-18 15:49:42

Update_time:NULL

Check_time:NULL

Collation:latin1_swedish_ci

Checksum:NULL

Create_options:row_format=REDUNDANT

Comment:

1 row in set(0.00 sec)

mysql□SELECT*FROM mytest2\G;

***************************1.row***************************

t1:a

t2:bb

t3:bb
```

```
t4:ccc

***************************2.row***************************

t1:d

t2:ee

t3:ee

t4:fff

***************************3.row***************************

t1:d

t2:NULL

t3:NULL

t4:fff

3 rows in set(0.00 sec)
```

□□□□□□□row_format□□Redundant□□□□□hexdump□□□□mytest2.ibd□□□□□□□mytest2.txt□□□□□□□□□□□□□□

```
0000c070 08 03 00 00 73 75 70 72 65 6d 75 6d 00 23 20 16|....supremum.#.|

0000c080 14 13 0c 06 00 00 10 0f 00 ba 00 00 00 2b 68 0b|............+h.|

0000c090 00 00 00 00 06 53 80 00 00 00 32 01 10 61 62 62|.....S....2..abb|

0000c0a0 62 62 20 20 20 20 20 20 20 20 63 63 63 23 20 16|bb ccc#.|

0000c0b0 14 13 0c 06 00 00 18 0f 00 ea 00 00 00 2b 68 0c|............+h.|

0000c0c0 00 00 00 00 06 53 80 00 00 00 32 01 1e 64 65 65|.....S....2..dee|

0000c0d0 65 65 20 20 20 20 20 20 20 20 66 66 66 21 9e 94|ee fff!..|

0000c0e0 14 13 0c 06 00 00 20 0f 00 74 00 00 00 2b 68 0d|........t...+h.|

0000c0f0 00 00 00 00 06 53 80 00 00 00 32 01 2c 64 00 00|.....S....2.,d..|

0000c100 00 00 00 00 00 00 00 00 66 66 66 00 00 00 00 00|........fff.....|
```

□□□□□□□□□□□

```
23 20 16 14 13 0c 06/*□□□□□□□□□□*/
```

```
00 00 10 0f 00 ba/*Record Header□□□6□□□*/

00 00 00 2b 68 0b/*RowID*/

00 00 00 00 06 53/*TransactionID*/

80 00 00 00 32 01 10/*Roll Point*/

61/*□1□□'a'*/

62 62/*□2□□'bb'*/

62 62 20 20 20 20 20 20 20 20/*□3□□'bb'Char□□*/

63 63 63/*□4□□'ccc'*/
```

23 20 16 14 13 0c 06□□□06□0c□13□14□16□20□23□□□□□□□□□□6□□□□□□6□6+6=0x0C□□□□□□□□7□6+6+7=0x13□□□□□□□1□6+6+7+1=0x14□□□□□□□2□6+6+7+1+2=0x16□□□□□□□10□6+6+7+1+2+10=0x20□□□□□□□3□6+6+7+1+2+10+3=0x23□□

□□□□□□□□□□□Record Header□□□□□□48□□□□22□32□□□000000111□□□□□□□7□□□□□□□□□3□□□□□□□□33□□1□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□Redundant□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□NULL□□□□□□□□□□□□□

```
21 9e 94 14 13 0c 06/*□□□□□□□□□*/

00 00 20 0f 00 74/*Record Header□□□6□□*/

00 00 00 2b 68 0d/*RowID*/

00 00 00 00 06 53/*TransactionID*/

80 00 00 00 32 01 10/*Roll Point*/

64/*□1□□'d'*/

00 00 00 00 00 00 00 00 00 00/*□3□□NULL*/

66 66 66/*□4□□'fff'*/
```

□□□□□Compact□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□06 0c 13 14 94 9e 21□□4□□□□□□□□□5□NULL□□□□94□□□□6□CHAR□□□NULL□□9e□94+10=0x9e□□□□□21□□□14+3=0x21□□□□□□□□VARCHAR□□□NULL□□Redundant□□□□□□□□□□□□□□□□□□CHAR□□□NULL□□□□□□□□

□□□mytest2□□□□□Latin1□□□□□□□□□□1□□□□□□□□mytest2□□□□□□□utf8□□□□CHAR□□□□□□□□□□□□10□□□□□□10×3=30□□□□□□Redundant□□□□□□□CHAR□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

4.3.3

InnoDBBLOBLOBBLOBVARCHAR

VARCHARDBAMySQLVARCHAROracle VARCHAR24000SQL Server8000MySQLVARCHAR6553565535VARCHAR65535

```
mysqlCREATE TABLE test(

-a VARCHAR(65535)

-)CHARSET=latin1 ENGINE=InnoDB;

ERROR 1118(42000):Row size too large.The maximum row size for the used table type,not counting BLOBs,is 65535.You have to change
some columns to TEXT or BLOBs
```

InnoDB65535VARCHARVARCHAR65532

```
mysqlCREATE TABLE test(

-a VARCHAR(65532)

-)CHARSET=latin1 ENGINE=InnoDB;

Query OK,0 rows affected(0.15 sec)
```

SQL_MODEMySQLwarning

```
mysqlCREATE TABLE test(

-a VARCHAR(65535)

-)CHARSET=latin1 ENGINE=InnoDB;
```

```
Query OK,0 rows affected,1 warning(0.14 sec)

mysql□SHOW WARNINGS\G;

***************************1.row***************************

Level:Note

Code:1246

Message:Converting column'a'from VARCHAR to TEXT

1 row in set(0.00 sec)
```

warning[illegible]MySQL[illegible]VARCHAR[illegible]TEXT[illegible]test[illegible]

```
mysql□SHOW CREATE TABLE test\G;

***************************1.row***************************

Table:test

Create Table:CREATE TABLE'test'(

'a'mediumtext

)ENGINE=InnoDB DEFAULT CHARSET=utf8

1 row in set(0.00 sec)
```

[illegible]VARCHAR[illegible]65 532[illegible]latin1[illegible]GBK[illegible]UTF-8[illegible]

```
mysql□CREATE TABLE test(

-□a VARCHAR(65532)

-□)CHARSET=GBK ENGINE=InnoDB;

ERROR 1074(42000):Column length too big for column'a'(max=32767);use BLOB or TEXT instead

mysql□mysql□CREATE TABLE test(

-□a VARCHAR(65532)

-□)CHARSET=UTF8 ENGINE=InnoDB;

ERROR 1074(42000):Column length too big for column'a'(max=21845);use BLOB or TEXT instead
```

[illegible]VARCHAR[illegible]65532[illegible]max[illegible]VARCHAR[illegible]N[illegible]N[illegible]VARCHAR[illegible]65535[illegible]

[illegible]MySQL[illegible]65535[illegible]VARCHAR[illegible]

```
mysql□CREATE TABLE test2(

-□a VARCHAR(22000),

-□b VARCHAR(22000),

-□c VARCHAR(22000)

-□)CHARSET=latin1 ENGINE=InnoDB;

ERROR 1118(42000):Row size too large.The maximum row size for the used table type,not counting BLOBs,is 65535.You have to change some columns to TEXT or BLOBs
```

3[illegible]66000[illegible]InnoDB[illegible]65532[illegible]InnoDB[illegible]16KB[illegible]16384[illegible]65532[illegible]InnoDB[illegible]B-tree node[illegible]Uncompress BLOB[illegible]

```
mysql□CREATE TABLE t(

-□a VARCHAR(65532)

-□)ENGINE=InnoDB CHARSET=latin1;

Query OK,0 rows affected(0.15 sec)

mysql□INSERT INTO t

-□SELECT REPEAT('a',65532);

Query OK,1 row affected(0.08 sec)

Records:1 Duplicates:0 Warnings:0
```

[illegible]a[illegible]65 532[illegible]VARCHAR[illegible]t[illegible]a[illegible]65 532[illegible]py_innodb_page_info[illegible]

```
[root@nineyou0-43 mytest]#py_innodb_page_info.py-v t.ibd

page offset 00000000,page type□File Space Header□

page offset 00000001,page type□Insert Buffer Bitmap□

page offset 00000002,page type□File Segment inode□

page offset 00000003,page type□B-tree Node□,page level□0000□

page offset 00000004,page type□Uncompressed BLOB Page□

page offset 00000005,page type□Uncompressed BLOB Page□

page offset 00000006,page type□Uncompressed BLOB Page□

page offset 00000007,page type□Uncompressed BLOB Page□

Total number of page:8:

Insert Buffer Bitmap:1

Uncompressed BLOB Page:4

File Space Header:1

B-tree Node:1

File Segment inode:1
```

□□□□□□□□□□□□□□□□□□□□□B-tree Node□□□□4□□□□□□□□□□□□Uncompressed BLOB Page□□□□□□□□□□□□65532□□□□□□□□□□□□□□□□□BLOB□□□□□□□□□□□□□□□□□□□□□□□□□□hexdump□□□□□□□□□□□□□c000□□□□□

```
0000c000 67 ce fc 0b 00 00 00 03 ff ff ff ff ff ff ff ff|g...............|

0000c010 00 00 00 0a 6a d9 c0 89 45 bf 00 00 00 00 00 00|....j...E.......|

0000c020 00 00 00 00 00 c3 00 02 03 a7 80 03 00 00 00 00|................|

0000c030 00 80 00 05 00 00 00 01 00 00 00 00 00 00 00 00|................|

0000c040 00 00 00 00 00 00 00 00 01 a1 00 00 00 c3 00 00|................|

0000c050 00 02 00 f2 00 00 00 c3 00 00 00 02 00 32 01 00|............2..|

0000c060 02 00 1d 69 6e 66 69 6d 75 6d 00 02 00 0b 00 00|...infimum......|

0000c070 73 75 70 72 65 6d 75 6d 14 c3 00 00 00 10 ff f0|supremum........|

0000c080 00 00 00 b6 2b 00 00 00 00 51 4b 06 80 00 00 00|....+....QK.....|
0000c090 2d 01 10 61 61 61 61 61 61 61 61 61 61 61 61 61|-..aaaaaaaaaaaaa|
0000c0a0 61 61 61 61 61 61 61 61 61 61 61 61 61 61 61 61|aaaaaaaaaaaaaaaa|
0000c0b0 61 61 61 61 61 61 61 61 61 61 61 61 61 61 61 61|aaaaaaaaaaaaaaaa|
0000c0c0 61 61 61 61 61 61 61 61 61 61 61 61 61 61 61 61|aaaaaaaaaaaaaaaa|
0000c0d0 61 61 61 61 61 61 61 61 61 61 61 61 61 61 61 61|aaaaaaaaaaaaaaaa|
0000c0e0 61 61 61 61 61 61 61 61 61 61 61 61 61 61 61 61|aaaaaaaaaaaaaaaa|
0000c0f0 61 61 61 61 61 61 61 61 61 61 61 61 61 61 61 61|aaaaaaaaaaaaaaaa|
0000c100 61 61 61 61 61 61 61 61 61 61 61 61 61 61 61 61|aaaaaaaaaaaaaaaa|
0000c110 61 61 61 61 61 61 61 61 61 61 61 61 61 61 61 61|aaaaaaaaaaaaaaaa|
......
0000c390 61 61 61 00 00 00 c3 00 00 00 04 00 00 00 26 00|aaa...........&.|
0000c3a0 00 00 00 00 00 fc fc 00 00 00 00 00 00 00 00 00|................|
```

可以看到，从0x0000c093到0x0000c392数据页面其实只保存了VARCHAR（65532）的前768字节的前缀（prefix）数据（这里都是a），之后是偏移量，指向行溢出页，也就是前面用户看到的Uncompressed BLOB Page。因此，对于行溢出数据，其存放采用图4-4的方式。

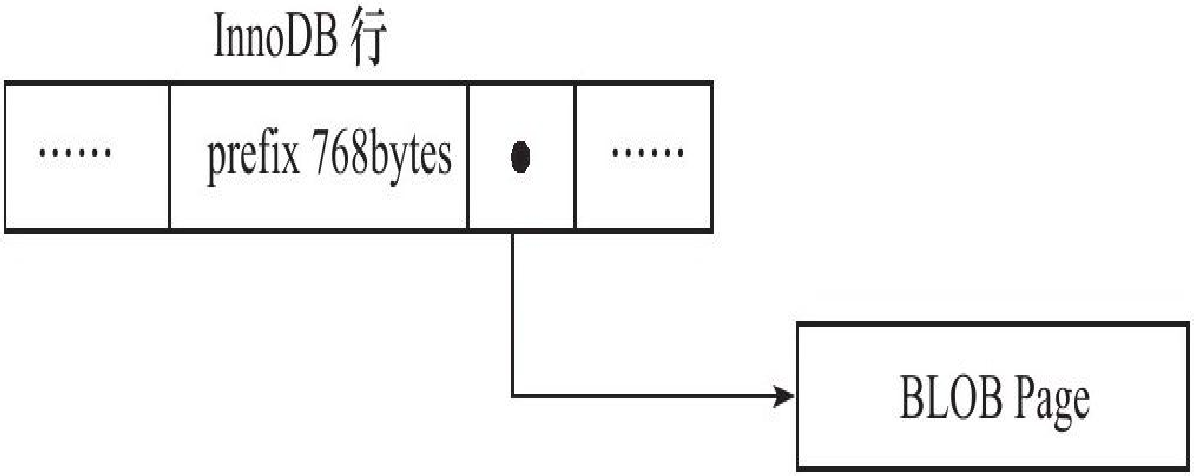


图 4-4 行溢出数据的存放

□□□□VARCHAR□□□□□□□□□□□□□□□□□□□□□□BLOB□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□B+Tree□□□□□□□□□□□□□□□□□□□□□□□□□□B+Tree□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
mysql□CREATE TALBE t(

-□a VARCHAR(9000)

-□)ENGINE=InnoDB;

Query OK,0 rows affected(0.13 sec)

mysql□INSERT INTO t

-□SELECT REPEAT('a',9000);

Query OK,1 row affected(0.04 sec)

Records:1 Duplicates:0 Warnings:0

mysql□INSERT INTO t

-□SELECT REPEAT('a',9000);

Query OK,1 row affected(0.04 sec)

Records:1 Duplicates:0 Warnings:0
```

□t□□□□□a□□□□9000□□□□□□□□□□□□□□□□□□□□□□□□□9000□□□□□□□□□□□□□□□□□□py_innodb_page_info□□□□□□□□□□□□□□□BLOB□□□

```
[root@nineyou0-43 mytest]#py_innodb_page_info.py-v t.ibd

page offset 00000000,page type□File Space Header□

page offset 00000001,page type□Insert Buffer Bitmap□

page offset 00000002,page type□File Segment inode□

page offset 00000003,page type□B-tree Node□,page level□0000□

page offset 00000004,page type□Uncompressed BLOB Page□

page offset 00000005,page type□Uncompressed BLOB Page□

Total number of page:6:

Insert Buffer Bitmap:1
```

```
Uncompressed BLOB Page:2

File Space Header:1

B-tree Node:1

File Segment inode:1
```

□□ □□py_innodb_page_info□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□VARCHAR□□□□□□□□□□□□BLOB□□□□□□□□□□□□□□□□□□□□□□□8098□□□□□□□□□□varchar□8098□□□□□□□□2□□□□

```
mysql□CREATE TABLE t(

-□a varchar(8098)

)ENGINE=InnoDB;

Query OK,0 rows affected(0.12 sec)

mysql□INSERT INTO t SELECT REPEAT('a',8098);

Query OK,1 row affected(0.04 sec)

Records:1 Duplicates:0 Warnings:0

mysql□INSERT INTO t SELECT REPEAT('a',8098);

Query OK,1 row affected(0.03 sec)

Records:1 Duplicates:0 Warnings:0
```

□□□py_innodb_page_info□□□□□□t.ibd□□□□□□□□□□□□□□□□□□□□□□□□□□□□□BLOB□□□□□□Microsoft SQL Server□□□□DBA□□□□□InnoDB□□□□□□VARCHAR□□□□□□SQL Server□□varchar□MAX□□□□□

```
[root@nineyou0-43 mytest]#py_innodb_page_info.py-v t.ibd

page offset 00000000,page type□File Space Header□

page offset 00000001,page type□Insert Buffer Bitmap□
page offset 00000002,page type□File Segment inode□

page offset 00000003,page type□B-tree Node□,page level□0000□

page offset 00000000,page type□Freshly Allocated Page□

page offset 00000000,page type□Freshly Allocated Page□

Total number of page:6:

Freshly Allocated Page:2

Insert Buffer Bitmap:1

File Space Header:1

B-tree Node:1

File Segment inode:1
```

□□□□□□□□□TEXT□BLOB□□□□□□□□□□□□□□□□□□Uncompressed BLOB Page□□□□□□□□□□□□□□□□□□□□□□BLOB□□□□□□□□□VARCHAR□□□□□□□□□□□□□□□□□□□□

```
mysql□CREATE TABLE t(

-□a BLOB

-□)ENGINE=InnoDB;

Query OK,0 rows affected(0.12 sec)

mysql□INSERT INTO t SELECT REPEAT('a',8000);

Query OK,1 row affected(0.03 sec)

Records:1 Duplicates:0 Warnings:0

mysql□INSERT INTO t SELECT REPEAT('a',8000);

Query OK,1 row affected(0.03 sec)

Records:1 Duplicates:0 Warnings:0

mysql□INSERT INTO t SELECT REPEAT('a',8000);

Query OK,1 row affected(0.01 sec)

Records:1 Duplicates:0 Warnings:0

mysql□INSERT INTO t SELECT REPEAT('a',8000);

Query OK,1 row affected(0.06 sec)
```

```
Records:1 Duplicates:0 Warnings:0
```

□□□□□□□□BLOB□□□□□□□□□□4□□□□□□8000□□□□□□□py_innodb_page_info□□□□□□t.ibd□□□□□□□□□□□□□□□□□□BLOB□□□

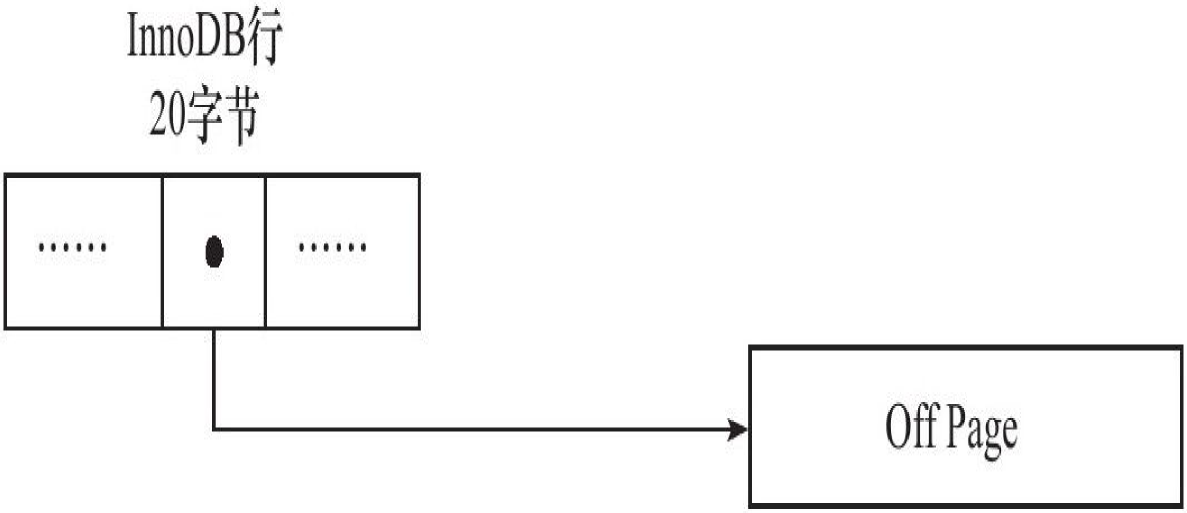
```
[root@nineyou0-43 mytest]#py_innodb_page_info.py-v t.ibd

page offset 00000000,page type□File Space Header□

page offset 00000001,page type□Insert Buffer Bitmap□

page offset 00000002,page type□File Segment inode□

page offset 00000003,page type□B-tree Node□,page level□0001□

page offset 00000004,page type□B-tree Node□,page level□0000□

page offset 00000005,page type□B-tree Node□,page level□0000□

page offset 00000006,page type□B-tree Node□,page level□0000□

page offset 00000000,page type□Freshly Allocated Page□

Total number of page:8:

Freshly Allocated Page:1

Insert Buffer Bitmap:1

File Space Header:1

B-tree Node:4

File Segment inode:1
```

□□□□□□□□□BLOB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□BLOB□□□□□□□□□□□□□□□□□□□□BLOB□□□□□□□□□□□□□768□□□

4.3.4 Compressed□Dynamic□□□□□

InnoDB 1.0.x□□□□□□□□□□□□□□file format□□□□□□□□□□□□□□□□□□□□Compact□Redundant□□□□Antelope□□□□□□□□□□□□□Barracuda□□□□□Barracuda□□□□□□□□□□□□□□□□□Compressed□Dynamic□

□□□□□□□□□□□□□BLOB□□□□□□□□□□□□□□□□□□□4-5□□□□□□□□□□□20□□□□□□□□□□□□□□□□Off Page□□□□□□Compact□Redundant□□□□□□□768□□□□□□



□ 4-5 Barracuda□□□□□□□□

Compressed□□□□□□□□□□□□□□□□□□□□□□□□□zlib□□□□□□□□□□□□BLOB□TEXT□VARCHAR□□□□□□□□□□□□□□□□□□□□□□

4.3.5 CHAR□□□□□□

□□□□VARCHAR□□□□□□□□□□□□□CHAR□□CHAR□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□latin1□□□□MySQL 4.1□□□□□CHR□N□□□N□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□CHAR□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
mysql□CREATE TABLE j(

-□a CHAR(2)

-□)CHARSET=GBK ENGINE=InnoDB;

Query OK,0 rows affected(0.11 sec)

mysql□INSERT INTO j SELECT'ab';

Query OK,1 row affected(0.03 sec)

Records:1 Duplicates:0 Warnings:0

mysql□SET NAMES GBK;

Query OK,0 rows affected(0.00 sec)

mysql□INSERT INTO j SELECT'□□';

Query OK,1 row affected(0.04 sec)

Records:1 Duplicates:0 Warnings:0

mysql□INSERT INTO j SELECT'a';

Query OK,1 row affected(0.03 sec)

Records:1 Duplicates:0 Warnings:0
```

□□□□□□□□j□□□□□GBK□□□□□□□□□□□□□□□'ab'□'□□'□□□□□□□□□□□□□□□□□

```
mysql□SELECT a,CHAR_LENGTH(a),LENGTH(a)

-□FROM j\G;
```

```
**************************1.row**************************
a:ab
CHAR_LENGTH(a):2
LENGTH(a):2
**************************2.row**************************
a:□□
CHAR_LENGTH(a):2
LENGTH(a):4
**************************3.row**************************
a:a
CHAR_LENGTH(a):1
LENGTH(a):1
3 rows in set(0.00 sec)
```

□□□□□CHAR_LENGTH□CHAR□□□□□□□□□□□□□'ab'□'□□'□□□□□□□□2□□□□□□□□'ab'□□2□□□□'□□'□□4□□□□□□□HEX□□□□□□□□□□□□□□□□□□□

```
mysql□SELECT a,HEX(a)
-□FROM j\G;
**************************1.row**************************
a:ab
HEX(a):6162
**************************2.row**************************
a:□□
HEX(a):CED2C3C7
**************************3.row**************************
a:a
HEX(a):61
3 rows in set(0.00 sec)
```

[illegible]'ab'[illegible]0x6162[illegible]'[illegible]'[illegible]0xCED2C3C7[illegible]CHAR[illegible]UTF-8[illegible]CHAR[illegible]10[illegible]10[illegible]30[illegible]CHAR[illegible]InnoDB[illegible]CHAR[illegible]hexdump[illegible]j.ibd[illegible]

```
0000c070 73 75 70 72 65 6d 75 6d 02 00 00 00 10 00 1c 00|supremum........|

0000c080 00 00 b6 2b 2b 00 00 00 51 52 da 80 00 00 00 2d|...++...QR.....-|

0000c090 01 10 61 62 04 00 00 00 18 ff d5 00 00 00 b6 2b|..ab...........+|

0000c0a0 2c 00 00 00 51 52 db 80 00 00 00 2d 01 10 ce d2|,...QR.....-....|

0000c0b0 c3 c7 00 00 00 00 00 00 00 00 00 00 00 00 00 00|................|
```

[illegible]

```
#[illegible]

02/*[illegible]2[illegible]CHAR[illegible]*/

00/*NULL[illegible]*/

00 00 10 00 1c/*Recoder Header*/

00 00 00 b6 2b 2b/*RowID*/

00 00 00 51 52 da/*TransactionID*/

80 00 00 00 2d 01 10/*Roll Point*/

61 62/*[illegible]'ab'*/

#[illegible]

04/*[illegible]4[illegible]CHAR[illegible]*/

00/*NULL[illegible]*/

00 00 18 ff d5/*Recoder Header*/

00 00 00 b6 2b 2c/*RowID*/

00 00 00 51 52 db/*TransactionID*/

80 00 00 00 2d 01 10/*Roll Point*/
```

```
c3 d2 c3 c7/*□□'□□'*/

#□□□□□

02/*□□□□□□2□□CHAR□□□□□□*/

00/*NULL□□□*/

00 00 20 ff b7/*Recoder Header*/

00 00 00 b6 2b 2d/*RowID*/

00 00 00 51 53 17/*TransactionID*/

80 00 00 00 2d 01 10/*Roll Point*/

61 20/*□□'a'*/
```

□□□□□□□□□□InnoDB□□□□□□□CHAR□□□□□□□□□□□□□□□CHAR□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□0x20□InnoDB□□□□□□□□□□□□□□□□□HEX□□□□□□□□□□□□□□□□□□□□□□□□□□□□□CHAR□VARCHAR□□□□□□□□□□□□□□□

4.4 InnoDB[illegible]

[illegible]InnoDB[illegible]B-tree Node[illegible]InnoDB[illegible]

[illegible] InnoDB[illegible]MySQL[illegible]InnoDB[illegible]InnoDB[illegible]Peter[illegible]InnoDB[illegible]Peter[illegible]InnoDB[illegible]Compact[illegible]

InnoDB[illegible]7[illegible]4-6[illegible]

❑File Header[illegible]

❑Page Header[illegible]

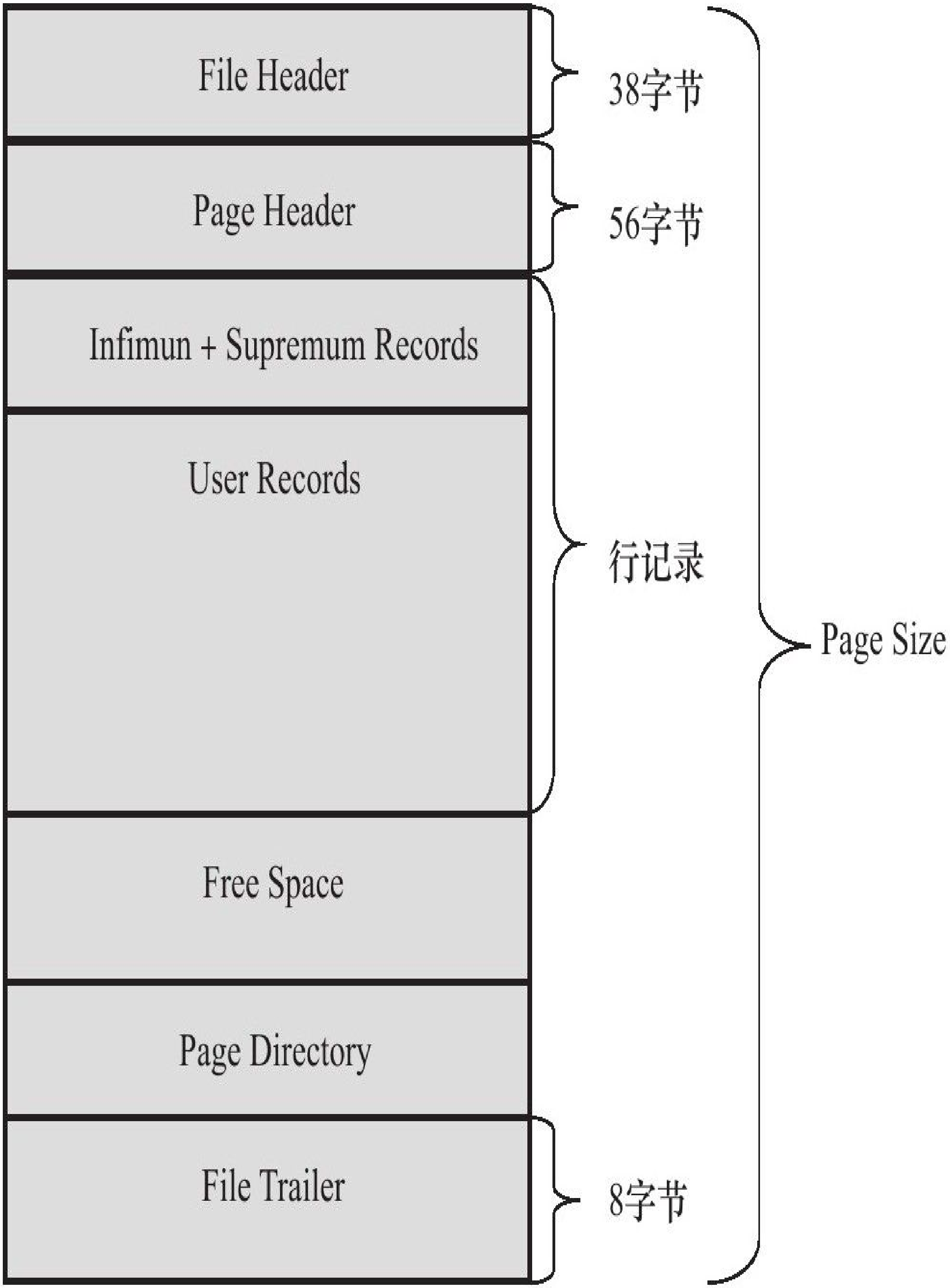
❑Infimun[illegible]Supremum Records

❑User Records[illegible]

❑Free Space[illegible]

❑Page Directory[illegible]

❑File Trailer[illegible]


File Header
38字节
Page Header
56字节
Infimun + Supremum Records
User Records
行记录
Page Size
Free Space
Page Directory
File Trailer
8字节

□ 4-6 InnoDB□□□□□□□□□

□□File Header□Page Header□File Trailer□□□□□□□□□□□38□56□8□□□□□□□□□□□□□□□□□□□□Checksum□□□□□□B+□□□□□□□□User Records□Free Space□Page Directory□□□

4.4.1 File Header

File Header□□□□□□□□□□□□□□4-3□8□□□□□□□□□38□□□

表4-3 File Header组成部分

名　称	大小（字节）	说　明
FIL_PAGE_SPACE_OR_CHKSUM	4	当MySQL为MySQL4.0.14之前的版本时，该值为0。在之后的MySQL版本中，该值代表页的checksum值（一种新的checksum值）
FIL_PAGE_OFFSET	4	表空间中页的偏移值。如某独立表空间a.ibd的大小为1GB，如果页的大小为16KB，那么总共有65 536个页。FIL_PAGE_OFFSET表示该页在所有页中的位置。若此表空间的ID为10，那么搜索页（10，1）就表示查找表a中的第二个页
FIL_PAGE_PREV	4	当前页的上一个页，B+ Tree特性决定了叶子节点必须是双向列表
FIL_PAGE_NEXT	4	当前页的下一个页，B+ Tree特性决定了叶子节点必须是双向列表
FIL_PAGE_LSN	8	该值代表该页最后被修改的日志序列位置LSN（Log Sequence Number）
FIL_PAGE_TYPE	2	InnoDB存储引擎页的类型。常见的类型见表4-4。记住0x45BF，该值代表了存放的是数据页，即实际行记录的存储空间
FIL_PAGE_FILE_FLUSH_LSN	8	该值仅在系统表空间的一个页中定义，代表文件至少被更新到了该LSN值。对于独立表空间，该值都为0
FIL_PAGE_ARCH_LOG_NO_OR_SPACE_ID	4	从MySQL 4.1开始，该值代表页属于哪个表空间

表 4-4 InnoDB 存储引擎中页的类型

名　称	十六进制	解　释
FIL_PAGE_INDEX	0x45BF	B+ 树叶节点
FIL_PAGE_UNDO_LOG	0x0002	Undo Log 页
FIL_PAGE_INODE	0x0003	索引节点
FIL_PAGE_IBUF_FREE_LIST	0x0004	Insert Buffer 空闲列表
FIL_PAGE_TYPE_ALLOCATED	0x0000	该页为最新分配
FIL_PAGE_IBUF_BITMAP	0x0005	Insert Buffer 位图
FIL_PAGE_TYPE_SYS	0x0006	系统页
FIL_PAGE_TYPE_TRX_SYS	0x0007	事务系统数据
FIL_PAGE_TYPE_FSP_HDR	0x0008	File Space Header
FIL_PAGE_TYPE_XDES	0x0009	扩展描述页
FIL_PAGE_TYPE_BLOB	0x000A	BLOB 页

4.4.2 Page Header

[illegible]File Header[illegible]Page Header[illegible]14[illegible]56[illegible]4-5[illegible]

表 4-5　Page Header 组成部分

名　　称	大小（字节）	说　　明
PAGE_N_DIR_SLOTS	2	在 Page Directory（页目录）中的 Slot（槽）数，“4.4.5 Page Directory” 小节中会介绍
PAGE_HEAP_TOP	2	堆中第一个记录的指针，记录在页中是根据堆的形式存放的
PAGE_N_HEAP	2	堆中的记录数。一共占用 2 字节，但是第 15 位表示行记录格式
PAGE_FREE	2	指向可重用空间的首指针
PAGE_GARBAGE	2	已删除记录的字节数，即行记录结构中 delete flag 为 1 的记录大小的总数
PAGE_LAST_INSERT	2	最后插入记录的位置
PAGE_DIRECTION	2	最后插入的方向。可能的取值为： ❑ PAGE_LEFT（0x01） ❑ PAGE_RIGHT（0x02） ❑ PAGE_SAME_REC（0x03） ❑ PAGE_SAME_PAGE（0x04） ❑ PAGE_NO_DIRECTION（0x05）
PAGE_N_DIRECTION	2	一个方向连续插入记录的数量
PAGE_N_RECS	2	该页中记录的数量
PAGE_MAX_TRX_ID	8	修改当前页的最大事务 ID，注意该值仅在 Secondary Index 中定义
PAGE_LEVEL	2	当前页在索引树中的位置，0x00 代表叶节点，即叶节点总是在第 0 层
PAGE_INDEX_ID	8	索引 ID，表示当前页属于哪个索引

（续）

名　称	大小（字节）	说　明
PAGE_BTR_SEG_LEAF	10	B+ 树数据页非叶节点所在段的 segment header。注意该值仅在 B+ 树的 Root 页中定义
PAGE_BTR_SEG_TOP	10	B+ 树数据页所在段的 segment header。注意该值仅在 B+ 树的 Root 页中定义

4.4.3 Infimum□Supremum Record

□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Infimum□□□□□□□□□□□□□□□□□□Supremum□□□Compact□□□□Redundant□□□□□□□□□□□□□□□□□□□4-7□□□Infimum□Supremum□□□

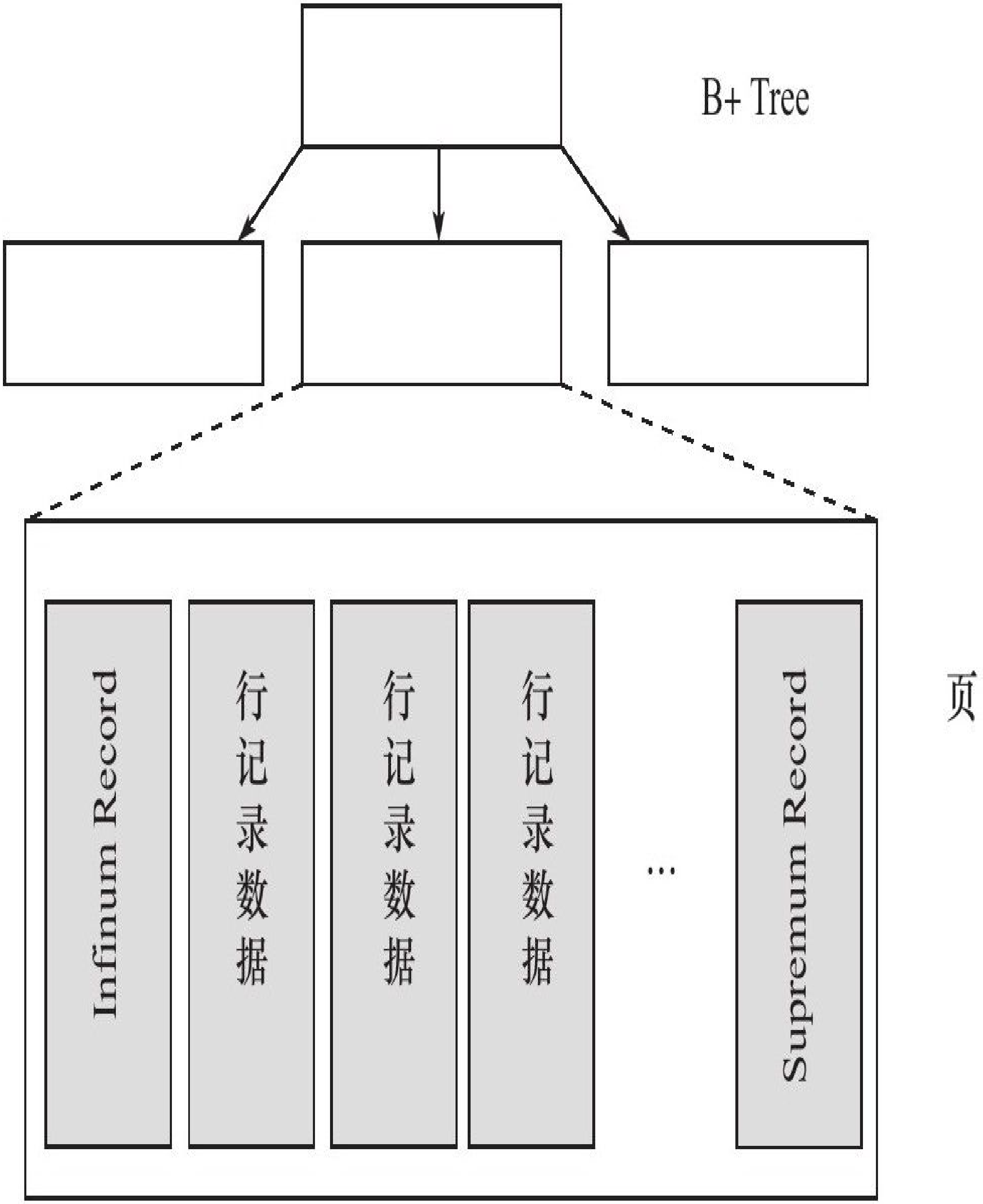


图 4-7 Infinum和Supremum Record

4.4.4 User Record[illegible]Free Space

User Record[illegible]InnoDB[illegible]B+[illegible]

Free Space[illegible]

4.4.5 Page Directory

Page Directory□□Slots□□□□□□□□Directory Slots□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□sparse directory□□□□□□□□□□□□□□□□□□□Infimum□n_owned□□□□1□□□Supremum□n_owned□□□□□□[1□8]□□□□□□□n_owned□□□□□□[4□8]□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□Slots□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□'i'□'d'□'c'□'b'□'e'□'g'□'l'□'h'□'f'□'j'□'k'□'a'□□□□□□□□□□□□4□□□□□Slots□□□□□□□□'a'□'e'□'i'□□

□□□InnoDB□□□□□Page Direcotry□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□recorder header□□next_record□□□□□□□□□□□□□Page Directory□□□□□□recorder header□□n_owned□□□□□□□□□□□□□□□□Page Directory□□

□□□□□□□B+□□Page Directory□□□

4.4.6 File Trailer

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□File Trailer□□□

File Trailer□□□□FIL_PAGE_END_LSN□□□□□8□□□□4□□□□□□□checksum□□□□4□□□File Header□□FIL_PAGE_LSN□□□□□□□□□File Header□□FIL_PAGE_SPACE_OR_CHKSUM□FIL_PAGE_LSN□□□□□□□□□□□□checksum□□□□□□□InnoDB□checksum□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□not corrupted□□

□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Corrupt□□□□□□File Trailer□□□□□□□□□□□□□□□□□□□□□□□□□□□□□innodb_checksums□□□□□□□□□□□□□□□□□

MySQL 5.6.6□□□□□□□□□innodb_checksum_algorithm□□□□□□□□□□checksum□□□□□□□□□□crc32□□□□□□□□innodb□crc32□none□strict_innodb□strict_crc32□strict_none□

innodb□□□□□□□InnoDB□□checksum□□□□□crc32□MySQL 5.6.6□□□□□□□checksum□□□□□□□□□□innodb□□□□□□□□□□□□□□□□□checksum□□□strict□□□□□□□□□□□MySQL□□□□□□□□□□□□none□□□□□□□checksum□□□

strict_*□□□□□□□□□□□□□□□checksum□□□□□□□□□□□□□□□MySQL□□□□□□MySQL 5.6.6□□□□□□□□□strict_crc32□□□□□□□□□□□□□strict_crc32□□□□□□□□□□□□□□innodb□crc32□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□mysql_upgrade□□□

4.4.7 InnoDB□□□□□□□□□

□□□□□□□□□□□□□□□□InnoDB□□t□□□□□□□□□□□

```
mysql□DROP TABLE IF EXISTS t;

Query OK,0 rows affected(0.04 sec)

mysql□CREATE TABLE t(

-□a INT UNSIGNED NOT NULL AUTO_INCREMENT,

-□b CHAR(10),

-□PRIMARY KEY(a),

-□)ENGINE=InnoDB CHARSET=UTF8;

Query OK,0 rows affected(0.00 sec)

mysql□DELIMITER$$

mysql□CREATE PROCEDURE load_t(count INT UNSIGNED)

-□BEGIN

-□SET@c=0;

-□WHILE@c□count DO

-□INSERT INTO t

-□SELECT NULL,REPEAT(CHAR(97+RAND()*26),10);

-□SET@c=@c+1;

-□END WHILE;

-□END;

-□$$

Query OK,0 rows affected(0.00 sec)

mysql□DELIMITER;

mysql□CALL load_t(100);

Query OK,0 rows affected(0.60 sec)
```

```
mysql□SELECTa,bFROM t LIMIT 10;

+----+---------------+

|a|b|

+----+---------------+

|1|dddddddddd|

|2|hhhhhhhhhh|

|3|bbbbbbbbbb|

|4|iiiiiiiiii|

|5|nnnnnnnnnn|

|6|qqqqqqqqqq|

|7|oooooooooo|

|8|yyyyyyyyyy|

|9|yyyyyyyyyy|

|10|vvvvvvvvvv|

+----+---------------+

10 rows in set(0.00 sec)
```

□□□□□py_innodb_page_info□□□t.ibd□□□□□□□:

```
[root@nineyou0-43 mytest]#py_innodb_page_info.py-v t.ibd

page offset 00000000,page type□File Space Header□

page offset 00000001,page type□Insert Buffer Bitmap□

page offset 00000002,page type□File Segment inode□

page offset 00000003,page type□B-tree Node□,page level□0000□

page offset 00000000,page type□Freshly Allocated Page□

page offset 00000000,page type□Freshly Allocated Page□

Total number of page:6:

Freshly Allocated Page:2

Insert Buffer Bitmap:1

File Space Header:1
```

```
B-tree Node:1

File Segment inode:1
```

[illegible]page offset 3[illegible]hexdump[illegible]t.ibd[illegible]0x0000c000[illegible]16K*3=0xc000[illegible]

```
0000c000 52 1b 24 00 00 00 00 03 ff ff ff ff ff ff ff ff|R.$.............|
0000c010 00 00 00 0a 6a e0 ac 93 45 bf 00 00 00 00 00 00|....j...E.......|
0000c020 00 00 00 00 00 dc 00 1a 0d c0 80 66 00 00 00 00|...........f....|
0000c030 0d a5 00 02 00 63 00 64 00 00 00 00 00 00 00 00|.....c.d........|
0000c040 00 00 00 00 00 00 00 00 01 ba 00 00 00 dc 00 00|................|
0000c050 00 02 00 f2 00 00 00 dc 00 00 00 02 00 32 01 00|.............2..|
0000c060 02 00 1c 69 6e 66 69 6d 75 6d 00 05 00 0b 00 00|...infimum......|
0000c070 73 75 70 72 65 6d 75 6d 0a 00 00 00 10 00 22 00|supremum......".|
0000c080 00 00 01 00 00 00 51 6d eb 80 00 00 00 2d 01 10|......Qm.....-..|
0000c090 64 64 64 64 64 64 64 64 64 64 0a 00 00 00 18 00|dddddddddd......|
0000c0a0 22 00 00 00 02 00 00 00 51 6d ec 80 00 00 00 2d|".......Qm.....-|
0000c0b0 01 10 68 68 68 68 68 68 68 68 68 68 0a 00 00 00|..hhhhhhhhhh....|
0000c0c0 20 00 22 00 00 00 03 00 00 00 51 6d ed 80 00 00|.".......Qm....|
0000c0d0 00 2d 01 10 62 62 62 62 62 62 62 62 62 62 0a 00|.-..bbbbbbbbbb..|
0000c0e0 04 00 28 00 22 00 00 00 04 00 00 00 51 6d ee 80|..(.".......Qm..|
0000c0f0 00 00 00 2d 01 10 69 69 69 69 69 69 69 69 69 69|...-..iiiiiiiiii|
0000c100 0a 00 00 00 30 00 22 00 00 00 05 00 00 00 51 6d|....0.".......Qm|
0000c110 ef 80 00 00 00 2d 01 10 6e 6e 6e 6e 6e 6e 6e 6e|.....-..nnnnnnnn|
0000c120 6e 6e 0a 00 00 00 38 00 22 00 00 00 06 00 00 00|nn....8.".......|
0000c130 51 6d f0 80 00 00 00 2d 01 10 71 71 71 71 71 71|Qm.....-..qqqqqq|
0000c140 71 71 71 71 0a 00 00 00 40 00 22 00 00 00 07 00|qqqq....@.".....|
0000c150 00 00 51 6d f1 80 00 00 00 2d 01 10 6f 6f 6f 6f|..Qm.....-..oooo|
0000c160 6f 6f 6f 6f 6f 6f 0a 00 04 00 48 00 22 00 00 00|oooooo....H."...|
0000c170 08 00 00 00 51 6d f2 80 00 00 00 2d 01 10 79 79|....Qm.....-..yy|

0000c180 79 79 79 79 79 79 79 79 0a 00 00 00 50 00 22 00|yyyyyyyy....P.".|

0000c190 00 00 09 00 00 00 51 6d f3 80 00 00 00 2d 01 10|......Qm.....-..|

0000c1a0 79 79 79 79 79 79 79 79 79 79 0a 00 00 00 58 00|yyyyyyyyyy....X.|

0000c1b0 22 00 00 00 0a 00 00 00 51 6d f4 80 00 00 00 2d|".......Qm.....-|

0000c1c0 01 10 76 76 76 76 76 76 76 76 76 76 0a 00 00 00|..vvvvvvvvvv....|

0000c1d0 60 00 22 00 00 00 0b 00 00 00 51 6d f5 80 00 00|'."......Qm....|

0000c1e0 00 2d 01 10 6b 6b 6b 6b 6b 6b 6b 6b 6b 6b 0a 00|.-..kkkkkkkkkk..|

0000c1f0 04 00 68 00 22 00 00 00 0c 00 00 00 51 6d f6 80|..h.".......Qm..|

……

0000ffc0 00 00 00 00 00 70 0d 1d 0c 95 0c 0d 0b 85 0a fd|.....p..........|

0000ffd0 0a 75 09 ed 09 65 08 dd 08 55 07 cd 07 45 06 bd|.u...e...U...E..|

0000ffe0 06 35 05 ad 05 25 04 9d 04 15 03 8d 03 05 02 7d|.5...%.........}|

0000fff0 01 f5 01 6d 00 e5 00 63 95 ae 5d 39 6a e0 ac 93|...m...c..]9j...|
```

[illegible]File Header[illegible]38[illegible]

- ❑52 1b 24 00[illegible]Checksum[illegible]
- ❑00 00 00 03[illegible]0[illegible]
- ❑ff ff ff ff[illegible]0xffffffff[illegible]
- ❑ff ff ff ff[illegible]0xffffffff[illegible]
- ❑00 00 00 0a 6a e0 ac 93[illegible]LSN[illegible]
- ❑45 bf[illegible]0x45bf[illegible]
- ❑00 00 00 00 00 00 00[illegible]
- ❑00 00 00 dc[illegible]SPACE ID[illegible]

□□□□□□□Page Header□□□□□□□File Trailer□□□□□File Trailer□□□□File Header□□□□□□□□□□□□□File Trailer□8□□□□

```
95 ae 5d 39 6a e0 ac 93
```

❑95 ae 5d 39□Checksum□□□□□□checksum□□□File Header□□□checksum□□□□□□

❑6a e0 ac 93□□□□□□File Header□□□□LSN□4□□□□□

□□□□56□□□Page Header□□□□□□□□□□□Page Header□□□□□□□□□□□□□□□□□□□□□□□□□

```
Page Header(56 bytes):

PAGE_N_DIR_SLOTS=0x001a

PAGE_HEAP_TOP=0x0dc0

PAGE_N_HEAP=0x8066

PAGE_FREE=0x0000

PAGE_GARBAGE=0x0000

PAGE_LAST_INSERT=0x0da5

PAGE_DIRECTION=0x0002

PAGE_N_DIRECTION=0x0063

PAGE_N_RECS=0x0064

PAGE_MAX_TRX_ID=0x0000000000000000

PAGE_LEVEL=00 00

PAGE_INDEX_ID=0x00000000000001ba

PAGE_BTR_SEG_LEAF=0x000000dc0000000200f2

PAGE_BTR_SEG_TOP=0x000000dc000000020032
```

PAGE_N_DIR_SLOTS=0x001a□□□Page Directory□26□□□□□□□□2□□□□□□□□0x0000ffc4□0x0000fff7□□□□□□□□

```
0000ffc0 00 00 00 00 00 70 0d 1d 0c 95 0c 0d 0b 85 0a fd|.....p..........|
0000ffd0 0a 75 09 ed 09 65 08 dd 08 55 07 cd 07 45 06 bd|.u...e...U...E..|
0000ffe0 06 35 05 ad 05 25 04 9d 04 15 03 8d 03 05 02 7d|.5...%.........}|
0000fff0 01 f5 01 6d 00 e5 00 63 95 ae 5d 39 6a e0 ac 93|...m...c..]9j...|
```

PAGE_HEAP_TOP=0x0dc0□□□□□□□□□□□□□□□□0xc000+0x0dc0=0xcdc0□□

```
0000cdb0 00 00 00 2d 01 10 70 70 70 70 70 70 70 70 70 70|...-..pppppppppp|
0000cdc0 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00|................|
0000cdd0 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00|................|
0000cde0 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00|................|
```

PAGE_N_HEAP=0x8066□□□□□□□□Compact□□□□□□0x0802□□□□□□Redundant□□□□□□□2□□□□□□□□□□□□□□□Infinimun□Supremum□□□□□□0x8066-0x8002=0x64□□□□□□□□□□□□□100□□□□

PAGE_FREE=0x0000□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□0□

PAGE_GARBAGE=0x0000□□□□□□□□□0□□□□□□□□□□□□□□□□□□□□□□□□□0□

PAGE_LAST_INSERT=0x0da5□□□□□□□□□□□□□□□□□□□□□□□□□□□0xc0000+0x0da5=0xcda5□□□□□□□

```
0000cda0 00 03 28 f2 cb 00 00 00 64 00 00 00 51 6e 4e 80|..(.....d...QnN.|
0000cdb0 00 00 00 2d 01 10 70 70 70 70 70 70 70 70 70 70|...-..pppppppppp|
0000cdc0 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00|................|
```

□□□□□□□□□□□a□□□100□□

PAGE_DIRECTION=0x0002□□□□□□□□□□□□□□□□□□□□□□PAGE_DIRECTION□□□□□□□□0x00002□

PAGE_N_DIRECTION=0x0063□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□100□□□□□□□□□99□

PAGE_N_RECS=0x0064□□□□□□□□□□□100□□□□□□PAGE_N_HEAP□□□□PAGE_N_HEAP□□□□□□□□□□□□□□□□□□□□□□□□□□□□0x8066□

PAGE_LEVEL=0x00□□□□□□□□□□□□□□□□□□□□□□□□□B+□□□□□□□□B+□□□□□□□0x00□

PAGE_INDEX_ID=0x00000000000001ba□□□ID□

□□□□□□□□Page Header□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
0000c050 00 02 00 f2 00 00 00 dc 00 00 00 02 00 32 01 00|.............2..|

0000c060 02 00 1c 69 6e 66 69 6d 75 6d 00 05 00 0b 00 00|...infimum......|

0000c070 73 75 70 72 65 6d 75 6d 0a 00 00 00 10 00 22 00|supremum......".|
```

□□0xc05E□0xc077□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□Char□8□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
#Infimum□□□□

01 00 02 00 1c/*recorder header*/

69 6e 66 69 6d 75 6d 00/*□□□□□□□□□□□□□□□□□Infimum□□□□□

/*0x00□□□*/

#Supremum□□□□

05 00 0b 00 00/*recorder header*/
```

```
73 75 70 72 65 6d 75 6d/*□□□□□□□□□□□□□□□□□Supremum*/
```

□□□□□infimum□□□□recorder header□□□□□□□□□□00 1c□□□□□□□□□□□□□□□□□□□□□□□□□□0xc063+0x001c□□0xc07f□0xc07f□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
0000c070 73 75 70 72 65 6d 75 6d 0a 00 00 00 10 00 22 00|supremum......".|

0000c080 00 00 01 00 00 00 51 6d eb 80 00 00 00 2d 01 10|......Qm.....-..|

0000c090 64 64 64 64 64 64 64 64 64 64 0a 00 00 00 18 00|dddddddddd......|

0000c0a0 22 00 00 00 02 00 00 00 51 6d ec 80 00 00 00 2d|".......Qm.....-|
```

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
/*□□□□□□*/

00 00 00 01/*□□□□□□□□□□□□□□□□□□□□ROWID□□a□□1*/

00 00 00 51 6d eb/*Transaction ID*/

80 00 00 00 2d 01 10/*Roll Pointer*/

64 64 64 64 64 64 64 64 64 64/*b□□□'aaaaaaaaaa'*/
```

□□□□□□□□□□□□□□

```
mysql□SELECT a,b,hex(b)FROM t ORDER BY a LIMIT 1;

+---+------------+-------------------------------+

|a|b|hex(b)|

+---+------------+-------------------------------+

|1|dddddddddd|64646464646464646464|

+---+------------+-------------------------------+

1 row in set(0.00 sec)
```

□□Recorder Header□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Page Header□PAGE_PREV□PAGE_NEXT□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□

□□□□Page Directory□□□□□□□□□0x0000ffc4□0x0000fff7□□□□□Page Directory□□□□

```
0000ffc0 00 00 00 00 00 70 0d 1d 0c 95 0c 0d 0b 85 0a fd|.....p..........|

0000ffd0 0a 75 09 ed 09 65 08 dd 08 55 07 cd 07 45 06 bd|.u...e...U...E..|

0000ffe0 06 35 05 ad 05 25 04 9d 04 15 03 8d 03 05 02 7d|.5...%.........}|

0000fff0 01 f5 01 6d 00 e5 00 63 95 ae 5d 39 6a e0 ac 93|...m...c..]9j...|
```

□□□□□□□Page Directory□□□□□□□□□□□□2□□□□□□□□00 63□□□□□□□□□□□0xc063□00 70□□□□□□□□□□□□□□□0xc070□□□□□□□□□□□□□Infimum□Supremum□□□□□□Page Directory□□□InnoDB□□□□□□□□□□□□□□□□Recorder Header□n_owned□□□□□□□□□□InnoDB□□□□□□□□□a□5□□□□□□□□□□Page Directory□□□□□□□□□□□□□□□00 e5□□□□□□□□□□□□0xc0e5□

```
0000c0e0 04 00 28 00 22 00 00 00 04 00 00 00 51 6d ee 80|..(."......Qm..|

0000c0f0 00 00 00 2d 01 10 69 69 69 69 69 69 69 69 69 69|...-..iiiiiiiii|

0000c100 0a 00 00 00 30 00 22 00 00 00 05 00 00 00 51 6d|....0.".......Qm|

0000c110 ef 80 00 00 00 2d 01 10 6e 6e 6e 6e 6e 6e 6e 6e|.....-..nnnnnnnn|

0000c120 6e 6e 0a 00 00 00 38 00 22 00 00 00 06 00 00 00|nn....8.".......|

0000c130 51 6d f0 80 00 00 00 2d 01 10 71 71 71 71 71 71|Qm.....-..qqqqqq|

0000c140 71 71 71 71 0a 00 00 00 40 00 22 00 00 00 07 00|qqqq....@.".....|
```

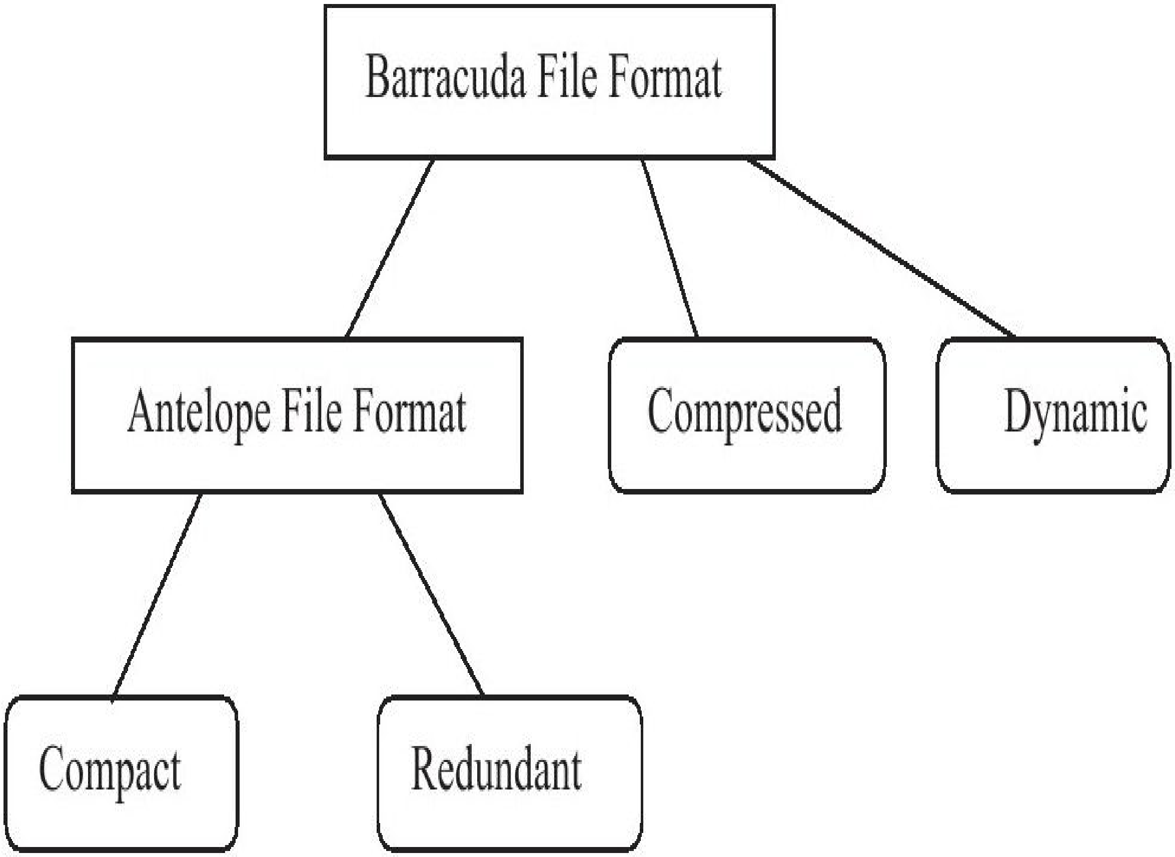
□□□□□□□□□□□4□□□□□□□□6□□□□□□□□□□5□□□Record Header□04 00 28 00 22□□□4□8□□□n_owned□□□□□□□□4□□□□□□□4□□□□□□□□□□□□□□□□□recorder header□□□□□□□□□□0x0022□□□□□□□□□□0xc107□□□□□□□□□□□□5□□□□

[illegible]InnoDB[illegible]

4.5 Named File Formats□□

□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB 1.0.x□□□□□□□□□□□□□□□□□□□□□□□□□□□□Off page□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB 1.0.x□□□□□InnoDB□□□□□Named File Formats□□□□□□□□□□□□□□□□□□□□□

InnoDB□□□□□1.0.x□□□□□□□□□□file format□□□□Antelope□□□□□□□□□□□□□□□□Barracuda□□□□□□□□□□□□□□□□□□□□□□□□4-8□□□Barracuda□□□□□Antelope□□□□□□□□□□Antelope□□□□□Compact□Redudant□□□□□Barracuda□□□□□□□□Antelope□□□□□□□□□□□□□□□□□□□□□□Compressed□Dynamic□□□□



□ 4-8 □□□□

InnoDB Plugin[illegible]InnoDB[illegible]Apple[illegible]

```
/**List of animal names representing file format.*/
static const char*file_format_name_map[]={
"Antelope",
"Barracuda",
"Cheetah",
"Dragon",
"Elk",
"Fox",
"Gazelle",
"Hornet",
"Impala",
"Jaguar",
"Kangaroo",
"Leopard",
"Moose",
"Nautilus",
"Ocelot",
"Porpoise",
"Quail",
"Rabbit",
"Shark",
"Tiger",
"Urchin",
"Viper",
"Whale",
"Xenops",
"Yak",

"Zebra"

};
```

□□innodb_file_format□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□

```
mysql□SELECT@@version\G;

***************************1.row***************************

@@version:5.1.37

1 row in set(0.00 sec)

mysql□SHOW VARIABLES LIKE'innodb_version'\G;

***************************1.row***************************

Variable_name:innodb_version

Value:1.0.4

1 row in set(0.00 sec)

mysql□SHOW VARIABLES LIKE'innodb_file_format'\G;

***************************1.row***************************

Variable_name:innodb_file_format

Value:Barracuda

1 row in set(0.00 sec)
```

□□innodb_file_format_check□□□□□□InnoDB□□□□□□□□□□□□□□□□□□ON□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
InnoDB:Warning:the system tablespace is in a

file format that this version doesn't support
```

4.6 □□

4.6.1 □□□□□

□□□constraint□□

□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□Primary Key□Unique Key□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□

❑□□□□□□□□□□□□□□□□□□□□□□□

❑□□□Foreign Key□□□□

❑□□□□□□

❑□□□□□□DEFAULT□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□InnoDB□□□

□□InnoDB□□□□□□□□□□□□□□□□□□□

❑Primary Key

❑Unique Key

❑Foreign Key

❑Default

❑NOT NULL

4.6.2 [illegible]

[illegible]

❑[illegible]

❑[illegible]ALTER TABLE[illegible]

[illegible]Unique Key[illegible]CREATE UNIQUE INDEX[illegible]PRIMARY[illegible]Unique Key[illegible]Unique Key[illegible]Foreign Key[illegible]

```
mysql[illegible]CREATE TABLE u(

-[illegible]id INT,

-[illegible]name VARCHAR(20),

-[illegible]id_card CHAR(18),

-[illegible]PRIMARY KEY(id),

-[illegible]UNIQUE KEY(name));

Query OK,0 rows affected(0.16 sec)

mysql[illegible]SELECT constraint_name,constraint_type

-[illegible]FROM

-[illegible]information_schema.TABLE_CONSTRAINTS

-[illegible]WHERE table_schema='mytest'AND table_name='u';\G;

***************************1.row***************************

constraint_name:PRIMARY

constraint_type:PRIMARY KEY

***************************2.row***************************

constraint_name:name

constraint_type:UNIQUE

2 rows in set(0.00 sec)
```

PRIMARYALTER TABLE:

```
mysqlALTER TABLE u
-ADD UNIQUE KEY uk_id_card(id_card);
Query OK,0 rows affected(0.19 sec)
Records:0 Duplicates:0 Warnings:0
mysqlSELECT constraint_name,constraint_type
-FROM
-information_schema.TABLE_CONSTRAINTS
-WHERE table_schema='mytest'AND table_name='u';\G;
***************************1.row***************************
constraint_name:PRIMARY
constraint_type:PRIMARY KEY
***************************2.row***************************
constraint_name:name
constraint_type:UNIQUE
***************************3.row***************************
constraint_name:uk_id_card
constraint_type:UNIQUE
3 rows in set(0.00 sec)
```

Foreign KeyForeign Keyp

```
mysqlCREATE TABLE p(
-id INT,
-u_id INT,
-PRIMARY KEY(id),
-□FOREIGN KEY(u_id)REFERENCES p(id));

Query OK,0 rows affected(0.13 sec)

mysql□SELECT constraint_name,constraint_type

-□FROM

-□information_schema.TABLE_CONSTRAINTS

-□WHERE table_schema='mytest'and table_name='p'\G;

***************************1.row***************************

constraint_name:PRIMARY

constraint_type:PRIMARY KEY

***************************2.row***************************

constraint_name:p_ibfk_1

constraint_type:FOREIGN KEY

2 rows in set(0.00 sec)
```

□□□□□□□□□□information_schema□□□□□TABLE_CONSTRAINTS□□□□□MySQL□□□□□□□□□□□□Foreign Key□□□□□□□□□□□□□□□□□REFERENTIAL_CONSTRAINTS□□□□□□□□□□□□□□□□□□

```
mysql□SELECT*FROM

-□information_schema.REFERENTIAL_CONSTRAINTS

-□WHERE constraint_schema='mytest'\G;

***************************1.row***************************

CONSTRAINT_CATALOG:NULL

CONSTRAINT_SCHEMA:test2

CONSTRAINT_NAME:p_ibfk_1

UNIQUE_CONSTRAINT_CATALOG:NULL

UNIQUE_CONSTRAINT_SCHEMA:test2

UNIQUE_CONSTRAINT_NAME:PRIMARY

MATCH_OPTION:NONE

UPDATE_RULE:RESTRICT
```

```
DELETE_RULE:RESTRICT

TABLE_NAME:p

REFERENCED_TABLE_NAME:p

1 row in set(0.00 sec)
```

4.6.3 [illegible]

[illegible]Primary Key[illegible]Unique Key[illegible]

[illegible]

4.6.4 对错误数据的约束

在某些默认设置下，MySQL数据库允许非法的或不正确的数据的插入或更新，又或者可以在数据库内部将其转化为一个合法的值，如向NOT NULL的字段插入一个NULL值，MySQL数据库会将其更改为0再进行插入，因此数据库本身没有对数据的正确性进行约束。例如：

```
mysql＞CREATE TABLE a(

-＞id INT NOT NULL,

-＞date DATE NOT NULL);

Query OK,0 rows affected(0.13 sec)

mysql＞INSERT INTO a

-＞SELECT NULL,'2009-02-30';

Query OK,1 row affected,2 warnings(0.04 sec)

Records:1 Duplicates:0 Warnings:2

mysql＞SHOW WARNINGS\G;

***************************1.row***************************

Level:Warning

Code:1048

Message:Column'id'cannot be null

***************************2.row***************************

Level:Warning

Code:1265

Message:Data truncated for column'date'at row 1

2 rows in set(0.00 sec)

mysql＞SELECT*FROM a\G;

***************************1.row***************************

id:0

date:0000-00-00

1 row in set(0.00 sec)
```

[illegible]NOT NULL[illegible]NULL[illegible]date[illegible][illegible]'2009-02-30'[illegible]“[illegible]”[illegible]MySQL[illegible]warning[illegible]MySQL[illegible]sql_mode[illegible]

```
mysql＞SET sql_mode='STRICT_TRANS_TABLES';

Query OK,0 rows affected(0.00 sec)

mysql＞INSERT INTO a

-＞SELECT NULL,'2009-02-30';

ERROR 1048(23000):Column'id'cannot be null

mysql＞INSERT INTO a

-＞SELECT 1,'2009-02-30';

ERROR 1292(22007):Incorrect date value:'2009-02-30'for column'date'at row 1
```

[illegible]sql_mode[illegible]STRICT_TRANS_TABLES[illegible]MySQL[illegible]sql_mode[illegible]MySQL[illegible]

4.6.5 ENUM□SET□□

MySQL□□□□□□□□□CHECK□□□□□□□ENUM□SET□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□male□female□□□□□□□□□□□□□ENUM□□□□□□□□

```
mysql□CREATE TABLE a(

-□id INT,

-□sex ENUM('male','female'));

Query OK,0 rows affected(0.12 sec)

mysql□INSERT INTO a

-□SELECT 1,'female';

Query OK,1 row affected(0.03 sec)

Records:1 Duplicates:0 Warnings:0

mysql□INSERT INTO a

-□SELECT 2,'bi';

Query OK,1 row affected,1 warning(0.03 sec)

Records:1 Duplicates:0 Warnings:1
```

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□CHECK□□□□□□□□□□□□sql_mode□

```
mysql□SET sql_mode='STRICT_TRANS_TABLES';

Query OK,0 rows affected(0.00 sec)

mysql□INSERT INTO a

-□SELECT 2,'bi';

ERROR 1265(01000):Data truncated for column'sex'at row 1
```

[illegible]CHECK[illegible]ENUM[illegible]SET[illegible]

4.6.6 [illegible]

[illegible]

[illegible]INSERT[illegible]DELETE[illegible]UPDATE[illegible]SQL[illegible]MySQL 5.0[illegible]MySQL 5.1[illegible]

[illegible]CREATE TRIGGER[illegible]Super[illegible]MySQL[illegible]

```
CREATE

[DEFINER={user|CURRENT_USER}]

TRIGGER trigger_name BEFORE|AFTER INSERT|UPDATE|DELETE

ON tbl_name FOR EACH ROW trigger_stmt
```

[illegible]6[illegible]INSERT[illegible]UPDATE[illegible]DELETE[illegible]BEFORE[illegible]AFTER[illegible]BEFORE[illegible]AFTER[illegible]MySQL[illegible]FOR EACH ROW[illegible]DB2[illegible]FOR EACH STATEMENT[illegible]

[illegible]MySQL[illegible]CHECK[illegible]

[illegible]“[illegible]”[illegible]

```
mysql[illegible]CREATE TABLE usercash(

-[illegible]userid INT NOT NULL,

-[illegible]cash INT UNSIGNED NOT NULL);

Query OK,0 rows affected(0.11 sec)
```

```
mysql□INSERT INTO usercash

-□SELECT 1,1000;

Query OK,1 row affected(0.03 sec)

Records:1 Duplicates:0 Warnings:0

mysql□UPDATE usercash

-□SET cash=cash-(-20)WHERE userid=1;

Query OK,1 row affected(0.05 sec)

Rows matched:1 Changed:1 Warnings:0
```

[illegible]SQL[illegible]

```
mysql□CREATE TABLE usercash_err_log(

-□userid INT NOT NULL,

-□old_cash INT UNSIGNED NOT NULL,

-□new_cash INT UNSIGNED NOT NULL,

-□user VARCHAR(30),

-□time DATETIME);

Query OK,0 rows affected(0.13 sec)

mysql□DELIMITER$$

Query OK,0 rows affected(0.00 sec)

mysql□CREATE TRIGGER tgr_usercash_update BEFORE UPDATE ON usercash

-□FOR EACH ROW

-□BEGIN

-□IF new.cash-old.cash□0 THEN

-□INSERT INTO usercash_err_log

-□SELECT old.userid,old.cash,new.cash,USER(),NOW();

-□SET new.cash=old.cash;

-□END IF;
```

```
-□END;

-□$$

Query OK,0 rows affected(0.00 sec)

mysql□DELIMITER$$

Query OK,0 rows affected(0.00 sec)
```

□□□□□□□□□□□□usercash_err_log□□□□□□□□□□□□□□□□□□□□□□□□□□□□tgr_usercash_update□□□□□BEFORE□□cash□□□□□□□□□□□□□□□□□□□□□usercash_err_log□□□□□□□□SQL□□□

```
mysql□DELETE FROM usercash;

Query OK,1 row affected(0.02 sec)

mysql□INSERT INTO usercash

-□SELECT 1,1000;

Query OK,1 row affected(0.03 sec)

Records:1 Duplicates:0 Warnings:0

mysql□UPDATE usercash

-□SET cash=cash-(-20)

-□WHERE userid=1;

Query OK,0 rows affected(0.02 sec)

Rows matched:1 Changed:0 Warnings:0

mysql□SELECT*FROM usercash\G;

***************************1.row***************************

userid:1

cash:100

1 row in set(0.00 sec)

mysql□SELECT*FROM usercash_err_log\G;

***************************1.row***************************

userid:1
```

```
old_cash:1000

new_cash:1020

user:root@localhost

time:2009-11-06 11:49:49

Message:Column'id'cannot be null

1 row in set(0.00 sec)
```

[illegible]usercash_err_log[illegible]SQL[illegible]

4.6.7 [illegible]

[illegible]MySQL[illegible]MyISAM[illegible]InnoDB[illegible]

```
[CONSTRAINT[symbol]]FOREIGN KEY
[index_name](index_col_name,...)
REFERENCES tbl_name(index_col_name,...)
[ON DELETE reference_option]
[ON UPDATE reference_option]
reference_option:
RESTRICT|CASCADE|SET NULL|NO ACTION
```

[illegible]CREATE TABLE[illegible]ALTER TABLE[illegible]

```
mysql[illegible]CREATE TABLE parent(
-[illegible]id INT NOT NULL,
-[illegible]PRIMARY KEY(id)
-[illegible])ENGINE=INNODB;
Query OK,0 rows affected(0.13 sec)
mysql[illegible]CREATE TABLE child(
-[illegible]id INT,parent_id INT,
-[illegible]FOREIGN KEY(parent_id)REFERENCES parent(id)
-[illegible])ENGINE=INNODB;
Query OK,0 rows affected(0.16 sec)
```

[illegible]ON DELETE[illegible]ON UPDATE[illegible]DELETE[illegible]UPDATE[illegible]

❑CASCADE

❑SET NULL

❑NO ACTION

❑RESTRICT

CASCADE[illegible]DELETE[illegible]UPDATE[illegible]DELETE[illegible]UPDATE[illegible]SET NULL[illegible]DELETE[illegible]UPDATE[illegible]NULL[illegible]NULL[illegible]NO ACTION[illegible]DELETE[illegible]UPDATE[illegible]RESTRICT[illegible]DELETE[illegible]UPDATE[illegible]ON DELETE[illegible]ON UPDATE[illegible]RESTRICT[illegible]

[illegible]Oracle[illegible]deferred check[illegible]SQL[illegible]MySQL[illegible]immediate check[illegible]MySQL[illegible]NO ACTION[illegible]RESTRICT[illegible]

[illegible]Oracle[illegible]InnoDB[illegible]Microsoft SQL Server[illegible]child[illegible]parent_id[illegible]SHOW CREATE TABLE[illegible]InnoDB[illegible]parent_id[illegible]

```
mysql[illegible]SHOW CREATE TABLE child\G;

***************************1.row***************************

Table:child

Create Table:CREATE TABLE'child'(

'id'int(11)DEFAULT NULL,

'parent_id'int(11)NOT NULL,

KEY'parent_id'('parent_id'),

CONSTRAINT'child_ibfk_1'FOREIGN KEY('parent_id')REFERENCES'parent'('id')

)ENGINE=InnoDB DEFAULT CHARSET=utf8
```

```
1 row in set(0.00 sec)
```

[illegible]MySQL[illegible]

```
mysql[illegible]SET foreign_key_checks=0;

mysql[illegible]LOAD DATA……

mysql[illegible]SET foreign_key_checks=1;
```

4.7 □□

□MySQL□□□□□□□□View□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□permanent table□□□□□□□□□□□□□□□□□□□□□□

4.7.1 □□□□□

□□base table□□

MySQL□□□□5.0□□□□□□□□□□□□□□□□□□□

```
CREATE

[OR REPLACE]

[ALGORITHM={UNDEFINED|MERGE|TEMPTABLE}]

[DEFINER={user|CURRENT_USER}]

[SQL SECURITY{DEFINER|INVOKER}]

VIEW view_name[(column_list)]

AS select_statement

[WITH[CASCADED|LOCAL]CHECK OPTION]
```

□□updatable view□□□□□□□□WITH CHECK OPTION□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□:

```
mysql□CREATE TABLE t(id INT);

Query OK,0 rows affected(0.13 sec)

mysql□CREATE VIEW v_t

-□AS

-□SELECT*FROM t WHERE id□10;
```

```
Query OK,0 rows affected(0.00 sec)

mysql[illegible]INSERT INTO v_t SELECT 20;

Query OK,1 row affected(0.03 sec)

Records:1 Duplicates:0 Warnings:0

mysql[illegible]SELECT*FROM v_t;

Empty set(0.00 sec)
```

[illegible]id[illegible]10[illegible]v_t[illegible]id[illegible]20[illegible]WITH CHECK OPTION[illegible]

```
mysql[illegible]ALTER VIEW v_t

-[illegible]AS

-[illegible]SELECT*FROM t WHERE id[illegible]10

-[illegible]WITH CHECK OPTION;

Query OK,0 rows affected(0.00 sec)

mysql[illegible]INSERT INTO v_t SELECT 20;

ERROR 1369(HY000):CHECK OPTION failed'mytest.v_t'
```

[illegible]MySQL[illegible]

MySQL[illegible]DBA[illegible]SHOW TABLES[illegible]

```
mysql[illegible]SHOW TABLES\G;

***************************1.row***************************

Tables_in_mytest:t

***************************2.row***************************

Tables_in_mytest[illegible]v_t

2 rows in set(0.00 sec)
```

□□SHOW TABLES□□□□t□□□v_t□□□□□□□□□□□□□□□□□□□□□□□□□□□information_schema□□□□TABLE□□□□□□□□□□□□BASE TABLE□□□SQL□□□□□

```
mysql□SELECT*FROM information_schema.TABLES

-□WHERE table_type='BASE TABLE'

-□AND table_schema=database()\G;

***************************1.row***************************

TABLE_CATALOG:NULL

TABLE_SCHEMA:mytest

TABLE_NAME:t

TABLE_TYPE:BASE TABLE

ENGINE:InnoDB

VERSION:10

ROW_FORMAT:Compact

TABLE_ROWS:1

AVG_ROW_LENGTH:16384

DATA_LENGTH:16384

MAX_DATA_LENGTH:0

INDEX_LENGTH:0

DATA_FREE:0

AUTO_INCREMENT:NULL

CREATE_TIME:2009-11-09 16:27:52

UPDATE_TIME:NULL

CHECK_TIME:NULL

TABLE_COLLATION:utf8_general_ci

CHECKSUM:NULL

CREATE_OPTIONS:

TABLE_COMMENT:
```

```
1 row in set(0.00 sec)
```

□□□□□□□□□□□□□meta data□□□□□□information_schema□□□□VIEWS□□□□□□□□□□□□□□□□□□□□□□□□definer□□□□□□□□□□□□□□□□□□□□□□□□VIEWS□□□□

```
mysql□SELECT*FROM

-□information_schema.VIEWS

-□WHERE table_schema=database()\G;

***************************1.row***************************

TABLE_CATALOG:NULL

TABLE_SCHEMA:mytest

TABLE_NAME:v_t

VIEW_DEFINITION:select'mytest'.'t'.'id'AS'id'from'mytest'.'t'where('mytest'.'t'.'id'□10)

CHECK_OPTION:CASCADED

IS_UPDATABLE:YES

DEFINER:root@localhost

SECURITY_TYPE:DEFINER

CHARACTER_SET_CLIENT:latin1

COLLATION_CONNECTION:latin1_swedish_ci

1 row in set(0.00 sec)
```

4.7.2 [illegible]

Oracle[illegible]——[illegible]JOIN[illegible]GROUP BY[illegible]SQL[illegible]Microsoft SQL Server[illegible]

[illegible]Oracle[illegible]

❑BUILD IMMEDIATE

❑BUILD DEFERRED

BUILD IMMEDIATE[illegible]BUILD DEFERRED[illegible]

[illegible]SQL[illegible]

[illegible]DML[illegible]

❑ON DEMAND

❑ON COMMIT

ON DEMAND[illegible]ON COMMIT[illegible]DML[illegible]

[illegible]

❑FAST

❑COMPLETE

❑FORCE

❑NEVER

FAST□□□□□□□□□□□□□□□□□□□□□□□□□COMPLETE□□□□□□□□□□□□□□□□□□□□□□FORCE□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□FAST□□□□□□□COMPLETE□□□□NEVER□□□□□□□□□□□□□□

MySQL□□□□□□□□□□□□□□□□□□□MySQL□□□ON DEMAND□□□

```
mysql□CREATE TABLE Orders
-□(
-□order_id INT UNSIGNED NOT NULL AUTO_INCREMENT,
-□product_name VARCHAR(30)NOT NULL,
-□price DECIMAL(8,2)NOT NULL,
-□amount SMALLINT NOT NULL,
-□PRIMARY KEY(order_id)
-□)ENGINE=InnoDB;
Query OK,0 rows affected(0.13 sec)
mysql□INSERT INTO Orders VALUES
-□(NULL,'CPU',135.5,1),
-□(NULL,'Memory',48.2,3),
-□(NULL,'CPU',125.6,3),
-□(NULL,'CPU',105.3,4)
-□;
Query OK,4 rows affected(0.03 sec)
Records:4 Duplicates:0 Warnings:0
mysql□SELECT*FROM Orders\G;
***************************1.row***************************
order_id:1
product_name:CPU
```

```
price:135.50
amount:1
***************************2.row***************************
order_id:2
product_name:Memory
price:48.20
amount:3
***************************3.row***************************
order_id:3
product_name:CPU
price:125.60
amount:3
***************************4.row***************************
order_id:4
product_name:CPU
price:105.30
amount:4
4 rows in set(0.00 sec)
```

□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
mysql□CREATE TABLE Orders_MV(
-□product_name VARCHAR(30)NOT NULL
-□,price_sum DECIMAL(8,2)NOT NULL
-□,amount_sum INT NOT NULL
-□,price_avg FLOAT NOT NULL
-□,orders_cnt INT NOT NULL
-□,UNIQUE INDEX(product_name)
-□);
Query OK,0 rows affected(0.13 sec)
```

```
mysql□INSERT INTO Orders_MV

-□SELECT product_name

-□,SUM(price),SUM(amount),AVG(price)

-□,COUNT(*)

-□FROM Orders

-□GROUP BY product_name;

Query OK,2 rows affected(0.02 sec)

Records:2 Duplicates:0 Warnings:0

mysql□

mysql□

mysql□SELECT*FROM Orders_MV\G;

***************************1.row***************************

product_name:CPU

price:366.40

amount_sum:8

price_avg:122.133

orders_cnt:3

***************************2.row***************************

product_name:Memory

price:48.20

amount_sum:3

price_avg:48.2

orders_cnt:1

2 rows in set(0.00 sec)
```

□□□□□□□□□□□□□□□□□□□Orders_MV□□□□_MV□□□□□□□DBA□□ON DEMAND□□□□□□□□□□□□□□□□□□□□□□□□□□□COMPLETE□□□□□□□□□FAST□□□□□□□□□□□□□□□□□□□□□□□□□□□□order_id□□□□

□□□□□□□□ON COMMIT□□□□□□□□□□□□□□□□□□Oracle□□□□□□□□□□□□□□□□□□□□□MySQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Orders□□□□□□□□□□□□□

```
DELIMITER$$

CREATE TRIGGER tgr_Orders_insert

AFTER INSERT ON Orders

FOR EACH ROW

BEGIN

SET@old_price_sum=0;

SET@old_amount_sum=0;

SET@old_price_avg=0;

SET@old_orders_cnt=0;

SELECT IFNULL(price_sum,0),IFNULL(amount_sum,0),IFNULL(price_avg,0),IFNULL(orders_cnt,0)

FROM Orders_MV

WHERE product_name=NEW.product_name

INTO@old_price_sum,@old_amount_sum,@old_price_avg,@old_orders_cnt;

SET@new_price_sum=@old_price_sum+NEW.price;

SET@new_amount_sum=@old_amount_sum+NEW.amount;

SET@new_orders_cnt=@old_orders_cnt+1;

SET@new_price_avg=@new_price_sum/@new_orders_cnt;

REPLACE INTO Orders_MV

VALUES(NEW.product_name,@new_price_sum,@new_amount_sum,@new_price_avg,@new_orders_cnt);

END;

$$

DELIMITER;
```

□□□□□□□□□INSERT□□□□□□□INSERT□□□□□□□□□Orders_MV□□□□□□□□□□□□□□□□□□□□□□□□□□Orders_MV□□□□□

```
mysql[illegible]INSERT INTO Orders VALUES(NULL,'SSD',299,3);

Query OK,1 row affected,1 warning(0.03 sec)

mysql[illegible]INSERT INTO Orders VALUES(NULL,'Memory',47.9,5);

Query OK,1 row affected(0.03 sec)

mysql[illegible]SELECT*FROM Orders_MV\G;

***************************1.row***************************

product_name:CPU

price:366.40

amount_sum:8

price_avg:122.133

orders_cnt:3

***************************2.row***************************

product_name:Memory

price:96.10

amount_sum:8

price_avg:48.05

orders_cnt:2

***************************3.row***************************

product_name:SSD

price:299.00

amount_sum:3

price_avg:299

orders_cnt:1

3 rows in set(0.00 sec)
```

[illegible]Orders_MV[illegible]SQL[illegible]ON_COMMIT[illegible]Orders[illegible]UPDATE[illegible]DELETE[illegible]DELETE[illegible]UPDATE[illegible]

[illegible]MySQL[illegible]MySQL[illegible]Query Rewrite[illegible]

4.8 [illegible]

4.8.1 [illegible]

[illegible]InnoDB[illegible]MyISAM[illegible]NDB[illegible]CSV[illegible]FEDORATED[illegible]MERGE[illegible]

MySQL[illegible]5.1[illegible]

MySQL[illegible][1][illegible][2][illegible]MySQL[illegible]MySQL[illegible]

[illegible]

```
mysql[illegible]SHOW VARIABLES LIKE'%partition%'\G;

***************************1.row***************************

Variable_name:have_partitioning

Value:YES

1 row in set(0.00 sec)
```

[illegible]SHOW PLUGINS[illegible]

```
mysql[illegible]SHOW PLUGINS\G;

……

***************************2.row***************************

Name:partition

Status:ACTIVE

Type:STORAGE ENGINE
```

```
Library:NULL

License:GPL

……

9 rows in set(0.01 sec)
```

[illegible]DBA[illegible]SQL[illegible]OLTP[illegible]

[illegible]MySQL[illegible]

❑RANGE[illegible]MySQL 5.5[illegible]RANGE COLUMNS[illegible]

❑LIST[illegible]RANGE[illegible]LIST[illegible]MySQL 5.5[illegible]LIST COLUMNS[illegible]

❑HASH[illegible]

❑KEY[illegible]MySQL[illegible]

[illegible]SQL[illegible]

```
mysql[illegible]CREATE TABLE t1(

-[illegible]col1 INT NOT NULL,

-[illegible]col2 DATE NOT NULL,

-[illegible]col3 INT NOT NULL,

-[illegible]col4 INT NOT NULL,

-[illegible]UNIQUE KEY(col1,col2)

-[illegible])

-[illegible]PARTITION BY HASH(col3)

-[illegible]PARTITIONS 4;
```

```
ERROR 1503(HY000):A PRIMARY KEY must include all columns in the table's partitioning function
```

[illegible]NULL[illegible]

```
mysql□CREATE TABLE t1(

-□col1 INT NULL,

-□col2 DATE NULL,

-□col3 INT NULL,

-□col4 INT NULL,

-□UNIQUE KEY(col1,col2,col3,col4)

-□)

-□PARTITION BY HASH(col3)

-□PARTITIONS 4;

Query OK,0 rows affected(0.53 sec)
```

[illegible]SQL[illegible]

```
CREATE TABLE t1(

col1 INT NULL,

col2 DATE NULL,

col3 INT NULL,

col4 INT NULL

)engine=innodb

PARTITION BY HASH(col3)

PARTITIONS 4;

CREATE TABLE t1(

col1 INT NULL,

col2 DATE NULL,
col3 INT NULL,

col4 INT NULL,

key(col4)

)engine=innodb

PARTITION BY HASH(col3)

PARTITIONS 4;
```

---

[1] [illegible]

[2] [illegible]

4.8.2 □□□□

1.RANGE□□

□□□□□□□□□□□□□RANGE□□□□□□□□□□□□□□□□□□□CREATE TABLE □□□□□□□id□□□□□□□□□id□□10□□□□□□p0□□□□id□□□□10□□20□□□□□□p1□□□

```
CREATE TABLE t(

id INT

)ENGINE=INNDB

PARTITION BY RANGE(id)(

PARTITION p0 VALUES LESS THAN(10),

PARTITION p1 VALUES LESS THAN(20));
```

□□□□□□□□□□□□□□□□□□□□□□□□□□ibd□□□□□□□□□□□□□□□□□□□ibd□□□□□□□□□t#P#p0.ibd□t#P#p1.ibd□

```
mysql□system ls-lh/usr/local/mysql/data/test2/t*

-rw-rw----1 mysql mysql 8.4K 7□31 14:11/usr/local/mysql/data/test2/t.frm

-rw-rw----1 mysql mysql 28 7□31 14:11/usr/local/mysql/data/test2/t.par

-rw-rw----1 mysql mysql 96K 7□31 14:12/usr/local/mysql/data/test2/t#P#p0.ibd

-rw-rw----1 mysql mysql 96K 7□31 14:12/usr/local/mysql/data/test2/t#P#p1.ibd
```

□□□□□□□□□

```
mysql□INSERT INTO t SELECT 9;

Query OK,1 row affected(0.03 sec)

Records:1 Duplicates:0 Warnings:0

mysql□INSERT INTO tSELECT 10;
```

```
Query OK,1 row affected(0.03 sec)

Records:1 Duplicates:0 Warnings:0

mysql□INSERT INTO t SELECT 15;

Query OK,1 row affected(0.03 sec)

Records:1 Duplicates:0 Warnings:0
```

□□□t□□□id□□□□□□□□□□□□□□id□□□□□□□□□□□□□□□□□□□□□□□□□□ information_schema□□□□PARTITIONS□□□□□□□□□□□□□□□

```
mysql□SELECT*FROM information_schema.PARTITIONS

-□WHERE table_schema=database()AND table_name='t'\G;

***************************1.row***************************

TABLE_CATALOG:NULL

TABLE_SCHEMA:test2

TABLE_NAME:t

PARTITION_NAME:p0

SUBPARTITION_NAME:NULL

PARTITION_ORDINAL_POSITION:1

SUBPARTITION_ORDINAL_POSITION:NULL

PARTITION_METHOD:RANGE

SUBPARTITION_METHOD:NULL

PARTITION_EXPRESSION:id

SUBPARTITION_EXPRESSION:NULL

PARTITION_DESCRIPTION:10

TABLE_ROWS:1

AVG_ROW_LENGTH:16384

DATA_LENGTH:16384

MAX_DATA_LENGTH:NULL

INDEX_LENGTH:0

DATA_FREE:0
```

```
CREATE_TIME:NULL
UPDATE_TIME:NULL
CHECK_TIME:NULL
CHECKSUM:NULL
PARTITION_COMMENT:
NODEGROUP:default
TABLESPACE_NAME:NULL
***************************2.row***************************
TABLE_CATALOG:NULL
TABLE_SCHEMA:test2
TABLE_NAME:t
PARTITION_NAME:p1
SUBPARTITION_NAME:NULL
PARTITION_ORDINAL_POSITION:2
SUBPARTITION_ORDINAL_POSITION:NULL
PARTITION_METHOD:RANGE
SUBPARTITION_METHOD:NULL
PARTITION_EXPRESSION:id
SUBPARTITION_EXPRESSION:NULL
PARTITION_DESCRIPTION:20
TABLE_ROWS:2
AVG_ROW_LENGTH:8192
DATA_LENGTH:16384
MAX_DATA_LENGTH:NULL
INDEX_LENGTH:0
DATA_FREE:0
CREATE_TIME:NULL
UPDATE_TIME:NULL
CHECK_TIME:NULL
CHECKSUM:NULL
```

```
PARTITION_COMMENT:

NODEGROUP:default

TABLESPACE_NAME:NULL

2 rows in set(0.00 sec)
```

TABLE_ROWS□□□□□□□□□□□□□□□□□□□□□□□□□9□10□15□□□□□□□□□□□□□□□□p0□□1□□□□□□p1□□2□□□□PARTITION_METHOD□□□□□□□□□□□□□□RANGE□

□□□t□□□MySQL□□□□□□□□□□□□□□□□□□□□t□□□30□□□□

```
mysql□INSERT INTO t SELECT 30;

ERROR 1526(HY000):Table has no partition for value 30
```

□□□□□□□□□□□□□□□□□□MAXVALUE□□□□□MAXVALUE□□□□□□□□□□□□□□□□□20□□□MAXVALUE□□□□□p2□□□

```
mysql□ALTER TABLE t

-□ADD PARTITION(

-□partition p2 values less than maxvalue);

Query OK,0 rows affected(0.45 sec)

Records:0 Duplicates:0 Warnings:0

mysql□INSERT INTO t SELECT 30;

Query OK,1 row affected(0.03 sec)

Records:1 Duplicates:0 Warnings:0
```

RANGE□□□sales□

```
mysql□CREATE TABLE sales(
-□money INT UNSIGNED NOT NULL,

-□date DATETIME

-□)ENGINE=INNODB

-□PARTITION by RANGE(YEAR(date))(

-□PARTITION p2008 VALUE LESS THEN(2009),

-□PARTITION p2009 VALUE LESS THEN(2010),

-□PARTITION p2010 VALUE LESS THEN(2011)

-□);

Query OK,0 rows affected(0.34 sec)

mysql□INSERT INTO sales SELECT 100,'2008-01-01';

Query OK,1 row affected(0.03 sec)

Records:1 Duplicates:0 Warnings:0

mysql□INSERT INTO sales SELECT 100,'2008-02-01';

Query OK,1 row affected(0.03 sec)

Records:1 Duplicates:0 Warnings:0

mysql□INSERT INTO sales SELECT 200,'2008-01-02';

Query OK,1 row affected(0.04 sec)

Records:1 Duplicates:0 Warnings:0

mysql□INSERT INTO sales SELECT 100,'2009-03-01';

Query OK,1 row affected(0.03 sec)

Records:1 Duplicates:0 Warnings:0

mysql□INSERT INTO sales SELECT 200,'2010-03-01';

Query OK,1 row affected(0.03 sec)

Records:1 Duplicates:0 Warnings:0
```

□□□□□□□□□□□□□sales□□□□□□□□□□□□□□□2008□□□□□□□□□□□DELETE FROM sales WHERE date□='2008-01-01'and date□'2009-01-01'□□□□□2008□□□□□□□□□□□□

```
mysql□alter table sales drop partition p2008;
```

```
Query OK,0 rows affected(0.18 sec)

Records:0 Duplicates:0 Warnings:0
```

2008

```
mysqlEXPLAIN PARTITIONS

-SELECT*FROM sales

-WHERE date='2008-01-01'AND date='2008-12-31'\G;

***************************1.row***************************

id:1

select_type:SIMPLE

table:sales

partitions:p2008

type:ALL

possible_keys:NULL

key:NULL

key_len:NULL

ref:NULL

rows:5

Extra:Using where

1 row in set(0.00 sec)
```

EXPLAIN PARTITIONSQLp2008——Partition Pruning

```
mysqlEXPLAIN PARTITIOENS

-SELECT*FROM sales

-WHERE date='2008-01-01'AND date'2009-01-01'\G;
```

```
***************************1.row***************************

id:1

select_type:SIMPLE

table:sales

partitions:p2008,p2009

type:ALL

possible_keys:NULL

key:NULL

key_len:NULL

ref:NULL

rows:5

Extra:Using where

1 row in set(0.00 sec)
```

□□□□□□date□'2009-01-01'□□□date□='2008-12-31'□□□□□□□□□□p2008□p2009□□SQL□□□

□□sales□□

```
mysql□CREATE TABLE sales(

-□money INT UNSIGNED NOT NULL,

-□date DATETIME

-□)ENGINE=INNODB

-□PARTITION by RANGE(YEAR(date)*100+MONTH(date))(

-□PARTITION p201001 VALUES LESS THEN(201002),

-□PARTITION p201002 VALUES LESS THEN(201003),

-□PARTITION p201003 VALUES LESS THEN(201004)

-□);

Query OK,0 rows affected(0.37 sec)
```

□□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□□
□□□□□□□

```
mysql□EXPLAIN PARTITIONS

-□SELECT*FROM sales

-□WHERE date□='2010-01-01'AND date□='2010-01-31'\G;

***************************1.row***************************

id:1

select_type:SIMPLE

table:sales

partitions:p201001,p201002,p201003

type:ALL

possible_keys:NULL

key:NULL

key_len:NULL

ref:NULL

rows:4

Extra:Using where

1 row in set(0.00 sec)
```

□□□□□□□□□p201001□p201002□p201003□□□□□□□□□□□□□□□□□□□□RANGE□□□□□□□□□□□□YEAR()□TO_DAYS()□TO_SECONDS()□UNIX_TIMESTAMP()□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□TO_DAYS□□□

```
mysql□CREATE TABLE sales(

-□money INT UNSIGNED NOT NULL,

-□date DATETO,E

-□)ENGINE=INNODB

-□PARTITION by range(TO_DAYS(date))(
-□PARTITION p201001

-□VALUES LESS THEN(TO_DAYS('2010-02-01')),

-□PARTITION p201002

-□VALUES LESS THEN(TO_DAYS('2010-03-01')),

-□PARTITION p201003

-□VALUES LESS THEN(TO_DAYS('2010-04-01'))

-□);

Query OK,0 rows affected(0.36 sec)
```

[illegible]

```
mysql□EXPLAIN PATITIONS

-□SELECT*FROM sales

-□WHERE date□='2010-01-01'AND date□='2010-01-31'\G;

***************************1.row***************************

id:1

select_type:SIMPLE

table:sales

partitions:p201001

type:ALL

possible_keys:NULL

key:NULL

key_len:NULL

ref:NULL

rows:4

Extra:Using where

1 row in set(0.00 sec)
```

2.LIST[illegible]

LIST□□□RANGE□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
mysql□CREATE TABLE t(

-□a INT,

-□b INT)ENGINE=INNODB

-□PARTITION BY LIST(b)(

-□PARTITION p0 VALUES IN(1,3,5,7,9),

-□PARTITION p1 VALUES IN(0,2,4,6,8)

-□);

Query OK,0 rows affected(0.26 sec)
```

□□□RANGE□□□□□□VALUES LESS THAN□□□LIST□□□□VALUES IN□□□□□□□□□□□□□□□□□□□□□□□□□□t□□□□□□□□

```
mysql□INSERT INTO t SELECT 1,1;

Query OK,1 row affected(0.03 sec)

Records:1 Duplicates:0 Warnings:0

mysql□INSERT INTO t SELECT 1,2;

Query OK,1 row affected(0.03 sec)

Records:1 Duplicates:0 Warnings:0

mysql□INSERT INTO t SELECT 1,3;

Query OK,1 row affected(0.03 sec)

Records:1 Duplicates:0 Warnings:0

mysql□INSERT INTO t SELECT 1,4;

Query OK,1 row affected(0.03 sec)

Records:1 Duplicates:0 Warnings:0

mysql□SELECT table_name,partition_name,table_rows

-□FROM information_schema.PARTITIONS

-□WHERE table_name='t'AND table_schema=DATABASE()\G;

***************************1.row***************************
```

```
table_name:t

partition_name:p0

table_rows:2

***************************2.row***************************

table_name:t

partition_name:p1

table_rows:2

2 rows in set(0.00 sec)
```

□□□□□□□□□□□□□□□MySQL□□□□□□□□□□□

```
mysql□INSERT INTO t SELECT 1,10;

ERROR 1526(HY000):Table has no partition for value 10
```

□□□□□INSERT□□□□□□□□□□□□□□□□□□□□□□MyISAM□InnoDB□□□□□□□□□□□□MyISAM□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□MyISAM□□□□□□□□□□□

```
mysql□CRATE TABLE t(

-□a INT,

-□b INT)ENGINE=MyISAM

-□PARTITION BY LIST(b)(

-□PARTITION p0 VALUES IN(1,3,5,7,9),

-□PARTITION p1 VALUES IN(0,2,4,6,8)

-□);

Query OK,0 rows affected(0.05 sec)

mysql□INSERT INTO t VALUES(1,2),(2,4),(6,10),(5,3);

ERROR 1526(HY000):Table has no partition for value 10

mysql□SELECT*FROM t;

+------+------+
|a|b|

+------+------+

|1|2|

|2|4|

+------+------+

2 rows in set(0.00 sec)
```

[illegible]6[illegible]10[illegible]5[illegible]3[illegible]1[illegible]2[illegible]2[illegible]4[illegible]InnoDB[illegible]

```
mysql[illegible]TRUNCATE TABLE t;

Query OK,2 rows affected(0.00 sec)

mysql[illegible]ALTER TABLE t ENGINE=InnoDB;

Query OK,0 rows affected(0.25 sec)

Records:0 Duplicates:0 Warnings:0

mysql[illegible]INSERT INTO t VALUES(1,2),(2,4),(6,10),(5,3);

ERROR 1526(HY000):Table has no partition for value 10

mysql[illegible]SELECT*FROM t;

Empty set(0.00 sec)
```

[illegible]6[illegible]10[illegible]t[illegible]

3.HASH[illegible]

HASH[illegible]RANGE[illegible]LIST[illegible]HASH[illegible]MySQL[illegible]

[illegible]HASH[illegible]CREATE TABLE[illegible]“PARTITION BY HASH[illegible]expr[illegible]”[illegible]“expr”[illegible]

MySQL□□□□□□□□□□□□□□□□□□□□□□□□“PARTITIONS num”□□□□□num□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□PARTITIONS□□□□□□□□□□□□□□1□

□□□□□□□□□□HASH□□□□t□□□□□□□b□□□

```
CREATE TABLE t_hash(

a INT,

b DATETIME

)ENGINE=InnoDB

PARTITION BY HASH(YEAR(b))

PARTITIONS 4;
```

□□□□□□□b□2010-04-01□□□□□t_hash□□□□□□□□□□□□□□□□

```
MOD(YEAR('2010-04-01')□4)

=MOD(2010,4)

=2
```

□□□□□□□□□p2□□□□□□□□□□□□□□□

```
mysql□INSERT INTO t_hash SELECT 1,'2010-04-01';

Query OK,1 row affected(0.04 sec)

Records:1 Duplicates:0 Warnings:0

mysql□SELECT table_name,partition_name,table_rows

-□FROM information_schema.PARTITIONS

-□WHERE table_schema=DATABASE()AND table_name='t_hash'\G;

***************************1.row***************************

table_name:t_hash

partition_name:p0
```

```
table_rows:0

**************************2.row**************************

table_name:t_hash

partition_name:p1

table_rows:0

**************************3.row**************************

table_name:t_hash

partition_name:p2

table_rows:1

**************************4.row**************************

table_name:t_hash

partition_name:p3

table_rows:0

4 rows in set(0.00 sec)
```

[illegible]p2[illegible]1[illegible]YEAR[illegible]HASH[illegible]

MySQL[illegible]LINEAR HASH[illegible]HASH[illegible]HASH[illegible]LINEAR HASH[illegible]LINEAR HASH[illegible]t_linear_hash[illegible]t_hash[illegible]

```
CREATE TABLE t_linear_hash(

a INT,

b DATETIME

)ENGINE=InnoDB

PARTITION BY LINEAR HASH(YEAR(b))

PARTITIONS 4;
```

□□□□'2010-04-01'□□□□□□MySQL□□□□□□□□□□□□□□□□□□□

❑□□□□□□□4□□□□2□□□V□V=POWER(2□CEILING(LOG(2□num)))=4□

❑□□□□N=YEAR('2010-04-01')□(V-1)=2□

□□□□□□□P2□□□□□□□□□□□□HASH□□□□□□□□□□□□□□□□□□□□□

```
mysql＞INSERT INTO t_linear_hash SELECT 1,'2010-04-01';

Query OK,1 row affected(0.02 sec)

Records:1 Duplicates:0 Warnings:0

mysql＞SELECT table_name,partition_name,table_rows

-＞FROM information_schema.PARTITIONS

-＞WHERE table_schema=DATABASE()

-＞AND table_name='t_linear_hash'\G;

***************************1.row***************************

table_name:t_linear_hash

partition_name:p0

table_rows:0

***************************2.row***************************

table_name:t_linear_hash

partition_name:p1

table_rows:0

***************************3.row***************************

table_name:t_linear_hash

partition_name:p2

table_rows:1

***************************4.row***************************

table_name:t_linear_hash

partition_name:p3
```

```
table_rows:0
4 rows in set(0.01 sec)
```

LINEAR HASH□□HASH□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

4.KEY□□

KEY□□□HASH□□□□□□□□□□□HASH□□□□□□□□□□□□□□□□KEY□□□□MySQL□□□□□□□□□□□□□□□NDB Cluster□□□MySQL□□□□□MD5□□□□□□□□□□□□□□□MySQL□□□□□□□□□□□□□□□□□□□□□PASSWORD()□□□□□□□□□□

```
mysql□CREATE TABLE t_key(
-□a INT,
-□b DATETIME)ENGINE=InnoDB
-□PARTITION BY KEY(b)
-□PARTITIONS 4;
Query OK,0 rows affected(0.43 sec)
```

□KEY□□□□□□□□LINEAR□□HASH□□□□□□□□□□□□□□□□□□□□□2□□□powers-of-two□□□□□□□□□□□□□□□□□

5.COLUMNS□□

□□□□□□RANGE□LIST□HASH□KEY□□□□□□□□□□□□□□□□□□□□□□interger□□□□□□□□□□□□□□□□□□□□□□□□□□□YEAR()□TO_DAYS()□MONTH()□□□□MySQL5.5□□□□□□COLUMNS□□□□□□RANGE□□□LIST□□□□□□□□COLUMNS□□RANGECOLUMNS□□□□□□□□□□□□□□□

COLUMNS□□□□□□□□□□□□

❑[illegible]INT[illegible]SMALLINT[illegible]TINYINT[illegible]BIGINT[illegible]FLOAT[illegible]DECIMAL[illegible]

❑[illegible]DATE[illegible]DATETIME[illegible]

❑[illegible]CHAR[illegible]VARCHAR[illegible]BINARY[illegible]VARBINARY[illegible]BLOB[illegible]TEXT[illegible]

[illegible]YEAR()[illegible]TO_DAYS()[illegible]COLUMNS[illegible]

```
CREATE TABLE t_columns_range(

a INT,

b DATETIME

)ENGINE=INNODB

PARTITION BY RANGE COLUMNS(B)(

PARTITION p0 VALUES LESS THAN('2009-01-01'),

PARTITION p1 VALUES LESS THAN('2010-01-01')

);
```

[illegible]

```
CREATE TABLE customers_1(

first_name VARCHAR(25),

last_name VARCHAR(25),

street_1 VARCHAR(30),

street_2 VARCHAR(30),

city VARCHAR(15),

renewal DATE

)

PARTITION BY LIST COLUMNS(city)(
PARTITION pRegion_1
VALUES IN('Oskarshamn','Högsby','Mönsterås'),
PARTITION pRegion_2
VALUES IN('Vimmerby','Hultsfred','Västervik'),
PARTITION pRegion_3
VALUES IN('Nässjö','Eksjö','Vetlanda'),
PARTITION pRegion_4
VALUES IN('Uppvidinge','Alvesta','Växjo')
);
```

□□RANGE COLUMNS□□□□□□□□□□□□□□□□□

```
CREATE TABLE rcx(
a INT,
b INT,
c CHAR(3),
d INT
)Engine=InnoDB
PARTITION BY RANGE COLUMNS(a,d,c)(
PARTITION p0 VALUES LESS THAN(5,10,'ggg'),
PARTITION p1 VALUES LESS THAN(10,20,'mmmm'),
PARTITION p2 VALUES LESS THAN(15,30,'sss'),
PARTITION p3 VALUES LESS THAN(MAXVALUE,MAXVALUE,MAXVALUE)
);
```

MySQL5.5□□□□COLUMNS□□□□□□□□RANGE□LIST□□□□□□□□RANGE COLUMNS□LIST COLUMNS□□□□□□□□□□

4.8.3

subpartitioningcomposite partitioningMySQLRANGELISTHASHKEY

```
mysqlCREATE TABLE ts(a INT,b DATE)engine=innodb

-PARTITION BY RANGE(YEAR(b))

-SUBPARTITION BY HASH(TO_DAYS(b))

-SUBPARTITIONS 2(

-PARTITION p0 VALUES LESS THAN(1990),

-PARTITION p1 VALUES LESS THAN(2000),

-PARTITION p2 VALUES LESS THAN MAXVALUE

-);

Query OK,0 rows affected(0.01 sec)

mysqlsystem ls-lh/usr/local/mysql/data/test2/ts*

-rw-rw----1 mysql mysql 8.4K Aug 1 15:50/usr/local/mysql/data/test2/ts.frm

-rw-rw----1 mysql mysql 96 Aug 1 15:50/usr/local/mysql/data/test2/ts.par

-rw-rw----1 mysql mysql 96K Aug 1 15:50/usr/local/mysql/data/test2/ts#P#p0#SP#p0sp0.ibd

-rw-rw----1 mysql mysql 96K Aug 1 15:50/usr/local/mysql/data/test2/ts#P#p0#SP#p0sp1.ibd

-rw-rw----1 mysql mysql 96K Aug 1 15:50/usr/local/mysql/data/test2/ts#P#p1#SP#p1sp0.ibd

-rw-rw----1 mysql mysql 96K Aug 1 15:50/usr/local/mysql/data/test2/

ts#P#p1#SP#p1sp1.ibd

-rw-rw----1 mysql mysql 96K Aug 1 15:50/usr/local/mysql/data/test2/ts#P#p2#SP#p2sp0.ibd

-rw-rw----1 mysql mysql 96K Aug 1 15:50/usr/local/mysql/data/test2/ts#P#p2#SP#p2sp1.ibd
```

tsbRANGEHASH3×2=6SUBPARTITIONts

```
mysql＞CREATE TABLE ts(a INT,b DATE)
-＞PARTITION BY RANGE(YEAR(b))
-＞SUBPARTITION BY HASH(TO_DAYS(b))(
-＞PARTITION p0 VALUES LESS THAN(1990)(
-＞SUBPARTITION s0,
-＞SUBPARTITION s1
-＞),
-＞PARTITION p1 VALUES LESS THAN(2000)(
-＞SUBPARTITION s2,
-＞SUBPARTITION s3
-＞),
-＞PARTITION p2 VALUES LESS THAN MAXVALUE(
-＞SUBPARTITION s4,
-＞SUBPARTITION s5
-＞)
-＞);
Query OK,0 rows affected(0.00 sec)
```

[illegible]

❑[illegible]

❑[illegible]SUBPARTITION[illegible]

```
mysql＞CREATE TABLE ts(a INT,b DATE)
-＞PARTITION BY RANGE(YEAR(b))
-＞SUBPARTITION BY HASH(TO_DAYS(b))(
-＞PARTITION p0 VALUES LESS THAN(1990)(
-＞SUBPARTITION s0,
-＞SUBPARTITION s1
-□),
-□PARTITION p1 VALUES LESS THAN(2000),
-□PARTITION p2 VALUES LESS THAN MAXVALUE(
-□SUBPARTITION s2,
-□SUBPARTITION s3
-□)
-□);
ERROR 1064(42000):Wrong number of subpartitions defined,mismatch with previous setting near'
PARTITION p2 VALUES LESS THAN MAXVALUE(
SUBPARTITION s2,
SUBPARTITION s3
)
)'at line 8
```

❑[illegible]SUBPARTITION[illegible]

❑[illegible]

```
mysql□CREATE TABLE ts(a INT,b DATE)
-□PARTITION BY RANGE(YEAR(b))
-□SUBPARTITION BY HASH(TO_DAYS(b))(
-□PARTITION p0 VALUES LESS THAN(1990)(
-□SUBPARTITION s0,
-□SUBPARTITION s1
-□),
-□PARTITION p1 VALUES LESS THAN(2000)(
-□SUBPARTITION s0,
-□SUBPARTITION s1
-□),
-□PARTITION p2 VALUES LESS THAN MAXVALUE(
-□SUBPARTITION s0,

-□SUBPARTITION s1

-□)

-□);

ERROR 1517(HY000):Duplicate partition name s0
```

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□6□□□□□□□/disk0□/disk1□/disk2□□□□□□□□□□□□

```
mysql□CREATE TABLE ts(a INT,b DATE)ENGINE=MYISAM

-□PARTITION BY RANGE(YEAR(b))

-□SUBPARTITION BY HASH(TO_DAYS(b))(

-□PARTITION p0 VALUES LESS THAN(2000)(

-□SUBPARTITION s0

-□DATA DIRECTORY='/disk0/data'

-□INDEX DIRECTORY='/disk0/idx',

-□SUBPARTITION s1

-□DATA DIRECTORY='/disk1/data'

-□INDEX DIRECTORY='/disk1/idx'

-□),

-□PARTITION p1 VALUES LESS THAN(2010)(

-□SUBPARTITION s2

-□DATA DIRECTORY='/disk2/data'

-□INDEX DIRECTORY='/disk2/idx',

-□SUBPARTITION s3

-□DATA DIRECTORY='/disk3/data'

-□INDEX DIRECTORY='/disk3/idx'

-□),

-□PARTITION p2 VALUES LESS THAN MAXVALUE(

-□SUBPARTITION s4
->DATA DIRECTORY='/disk4/data'

->INDEX DIRECTORY='/disk4/idx',

->SUBPARTITION s5

->DATA DIRECTORY='/disk5/data'

->INDEX DIRECTORY='/disk5/idx'

->)

->);

Query OK,0 rows affected(0.02 sec)
```

□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□DATA DIRECTORY □INDEX DIRECTORY□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
mysql>CREATE TABLE ts(a INT,b DATE)engine=innodb

->PARTITION BY RANGE(YEAR(b))

->SUBPARTITION BY HASH(TO_DAYS(b))(

->PARTITION p0 VALUES LESS THAN(2000)(

->SUBPARTITION s0

->DATA DIRECTORY='/disk0/data'

->INDEX DIRECTORY='/disk0/idx',

->SUBPARTITION s1

->DATA DIRECTORY='/disk1/data'

->INDEX DIRECTORY='/disk1/idx'

->),

->PARTITION p1 VALUES LESS THAN(2010)(

->SUBPARTITION s2

->DATA DIRECTORY='/disk2/data'

->INDEX DIRECTORY='/disk2/idx',

->SUBPARTITION s3

->DATA DIRECTORY='/disk3/data'

->INDEX DIRECTORY='/disk3/idx'
-),

-PARTITION p2 VALUES LESS THAN MAXVALUE(

-SUBPARTITION s4

-DATA DIRECTORY='/disk4/data'

-INDEX DIRECTORY='/disk4/idx',

-SUBPARTITION s5

-DATA DIRECTORY='/disk5/data'

-INDEX DIRECTORY='/disk5/idx'

-)

-);

Query OK,0 rows affected(0.02 sec)

mysqlsystem ls-lh/usr/local/mysql/data/test2/ts*

-rw-rw----1 mysql mysql 8.4K Aug 1 16:24/usr/local/mysql/data/test2/ts.frm

-rw-rw----1 mysql mysql 80 Aug 1 16:24/usr/local/mysql/data/test2/ts.par

-rw-rw----1 mysql mysql 96K Aug 1 16:25/usr/local/mysql/data/test2/ts#P#p0#SP#s0.ibd

-rw-rw----1 mysql mysql 96K Aug 1 16:25/usr/local/mysql/data/test2/ts#P#p0#SP#s1.ibd

-rw-rw----1 mysql mysql 96K Aug 1 16:25/usr/local/mysql/data/test2/ts#P#p1#SP#s2.ibd

-rw-rw----1 mysql mysql 96K Aug 1 16:25/usr/local/mysql/data/test2/ts#P#p1#SP#s3.ibd

-rw-rw----1 mysql mysql 96K Aug 1 16:25/usr/local/mysql/data/test2/ts#P#p2#SP#s4.ibd

-rw-rw----1 mysql mysql 96K Aug 1 16:25/usr/local/mysql/data/test2/ts#P#p2#SP#s5.ibd
```

4.8.4 □□□□NULL□

MySQL□□□□□□NULL□□□□□□□□□□□□□□□□□□□□□□□□□MYSQL□□□□□□□□□NULL□□□□□□□□□□NULL□□□□MySQL□□□□□□NULL□□ORDER BY□□□□□□□□□□□□□□□□□□□MySQL□□□□□NULL□□□□□□□□□□□

□□RANGE□□□□□□□□□□□□NULL□□□MySQL□□□□□□□□□□□□□□□□□□□

```
mysql□CREATE TABLE t_range(

-□a INT,

-□b INT)ENGINE=InnoDB

-□PARTITION BY RANGE(b)(

-□PARTITION p0 VALUES LESS THAN(10),

-□PARTITION p1 VALUES LESS THAN(20),

-□PARTITION p2 VALUES LESS THAN MAXVALUE

-□);

Query OK,0 rows affected(0.01 sec)
```

□□□□□□□□1,1□□□1,NULL□□□□□□□□□□□□□□□□□□□□

```
mysql□INSERT INTO t_range SELECT 1,1;

Query OK,1 row affected(0.00 sec)

Records:1 Duplicates:0 Warnings:0

mysql□INSERT INTO t_range SELECT 1,NULL;

Query OK,1 row affected(0.00 sec)

Records:1 Duplicates:0 Warnings:0

mysql□SELECT*FROM t_range\G;

***************************1.row***************************
```

```
a:1

b:1

***************************2.row***************************

a:1

b:NULL

2 rows in set(0.00 sec)

mysql□SELECT table_name,partition_name,table_rows

-□FROM information_schema.PARTITIONS

-□WHERE table_schema=DATABASE()AND table_name='t_range'\G;

***************************1.row***************************

table_name:t_range

partition_name:p0

table_rows:2

***************************2.row***************************

table_name:t_range

partition_name:p1

table_rows:0

***************************3.row***************************

table_name:t_range

partition_name:p2

table_rows:0

3 rows in set(0.00 sec)
```

□□□□□□□□□□□□p0□□□□□□□□□RANGE□□□□NULL□□□□□□□□□□□□□□□□□□□□□□□□p0□□□□□□□□□□□□10□□□□□□□□NULL□□□□□□□□□□□□

```
mysql□ALTER TABLE t_range DROP PARTITION p0;

Query OK,0 rows affected(0.01 sec)

Records:0 Duplicates:0 Warnings:0
```

```
mysql□SELECT*FROM t_range;

Empty set(0.00 sec)
```

□LIST□□□□□□NULL□□□□□□□□□□□□□□□□□NULL□□□□□□□□□□

```
mysql□CREATE TABLE t_list(

-□a INT,

-□b INT)ENGINE=INNODB

-□PARTITION BY LIST(b)(

-□PARTITION p0 VALUES IN(1,3,5,7,9),

-□PARTITION p1 VALUES IN(0,2,4,6,8)

);

Query OK,0 rows affected(0.00 sec)

mysql□INSERT INTO t_list SELECT 1,NULL;

ERROR 1526(HY000):Table has no partition for value NULL
```

□p0□□□□NULL□□□□□□□□□□

```
mysql□CREATE TABLE t_list(

-□a INT,

-□b INT)ENGINE=INNODB

-□PARTITION BY LIST(b)(

-□PARTITION p0 VALUES IN(1,3,5,7,9,NULL),

-□PARTITION p1 VALUES IN(0,2,4,6,8)

);

Query OK,0 rows affected(0.00 sec)

mysql□INSERT INTO t_list SELECT 1,NULL;

Query OK,1 row affected(0.00 sec)

Records:1 Duplicates:0 Warnings:0

mysql□SELECT table_name,partition_name,table_rows
```

```
-□FROM information_schema.PARTITIONS

-□WHERE table_schema=DATABASE()AND table_name='t_list'\G;

***************************1.row***************************

table_name:t_list

partition_name:p0

table_rows:1

***************************2.row***************************

table_name:t_list

partition_name:p1

table_rows:0

2 rows in set(0.00 sec)
```

HASH□KEY□□□□NULL□□□□□□RANGE□□□LIST□□□□□□□□□□□□□□□□□□NULL□□□□□□□0□□□

```
mysql□CREATE TABLE t_hash(

-□a INT,

-□b INT)ENGINE=InnoDB

-□PARTITION BY HASH(b)

-□PARTITIONS 4;

Query OK,0 rows affected(0.00 sec)

mysql□INSERT INTO t_hash SELECT 1,0;

Query OK,1 row affected(0.00 sec)

Records:1 Duplicates:0 Warnings:0

mysql□INSERT INTO t_hash SELECT 1,NULL;

Query OK,1 row affected(0.01 sec)

Records:1 Duplicates:0 Warnings:0

mysql□SELECT table_name,partition_name,table_rows

-□FROM information_schema.PARTITIONS

-□WHERE table_schema=DATABASE()AND table_name='t_hash'\G;
```

```
***************************1.row***************************
table_name:t_hash
partition_name:p0
table_rows:2
***************************2.row***************************
table_name:t_hash
partition_name:p1
table_rows:0
***************************3.row***************************
table_name:t_hash
partition_name:p2
table_rows:0
***************************4.row***************************
table_name:t_hash
partition_name:p3
table_rows:0
4 rows in set(0.00 sec)
```

4.8.5 □□□□□

□□□□□□□□□“□□□□□□”□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ □□

□□□□□□□□□□□□□□OLTP□□□□□□□□□□Blog□□□□□□□□□□□□□□□□ OLAP□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□OLTP□□□□□□ OLAP□□□□□□□□□□□□□□□□□□□□□□□□□OLTP□□□□□□□□□□□□□□□□□ □□OLAP□□□□

□□OLAP□□□□□□□□□□□□□□□□□□□□□□□□OLAP□□□□□□□□□□□□□□□□ □□□□□□□□□□1□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Partition Pruning□□□

□□□□OLTP□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□10%□□□□□□ □□□□□□□□□□□□□□□□□□□B+□□□□□□□□□□□□□□□□□□□B+□□□2□3□ □□□IO□□□B+□□

□□□□□□□□□□□□□□1000W□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ □□10□HASH□□□□□□□□□□□□□100W□□□□□□□□□□□□□□□□□□ SELECT*FROM TABLE WHERE PK=@pk□□□□□□□□□□□□□□□□ 100W□1000W□□□□□□□□□B+□□□□□□□□□□□□□□2□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□1000W□B+□□□□□3□100W□B+□□□□□2□ □□□□□□□□□□□□□□□□1□IO□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□SQL□□□SELECT*FROM TABLE WHERE KEY=@key□□□□□KEY□□□□□□□□□□10□□□□□□□□□□□□□□□□2□IO□ □□□□□20□IO□□□□□□□□□□□□□□KEY□□□□□□2□3□IO□

□□□□□□□□Profile□□□□□ID□□□HASH□□□HASH□□□□□□10□□Profile □□□1000W□□□□

```
mysql□CREATE TABLE'Profile'(

-□'id'int(11)NOT NULL AUTO_INCREMENT,

-□'nickname'varchar(20)NOT NULL DEFAULT'',
-□'password'varchar(32)NOT NULL DEFAULT'',

-□'sex'char(1)NOT NULL DEFAULT'',

-□'rdate'date NOT NULL DEFAULT'0000-00-00',

-□PRIMARY KEY('id'),

-□KEY'nickname'('nickname')

-□)ENGINE=InnoDB

-□PARTITION BY HASH(id)

-□PARTITIONS 10;

Query OK,0 rows affected(1.29 sec)

mysql□SELECT COUNT(nickname)FROM Profile;

***************************1.row***************************

count(1):9999248

1 row in set(1 min 24.62 sec)
```

[illegible]HASH[illegible]

```
mysql□SELECT table_name,partition_name,table_rows

-□FROM information_schema.PARTITIONS

-□WHERE table_schema=DATABASE()AND table_name='Profile'\G;

***************************1.row***************************

table_name:Profile

partition_name:p0

table_rows:990703

***************************2.row***************************

table_name:Profile

partition_name:p1

table_rows:1086519

***************************3.row***************************

table_name:Profile

partition_name:p2
```

```
table_rows:976474
***************************4.row***************************
table_name:Profile
partition_name:p3
table_rows:986937
***************************5.row***************************
table_name:Profile
partition_name:p4
table_rows:993667
***************************6.row***************************
table_name:Profile
partition_name:p5
table_rows:978046
***************************7.row***************************
table_name:Profile
partition_name:p6
table_rows:990703
***************************8.row***************************
table_name:Profile
partition_name:p7
table_rows:978639
***************************9.row***************************
table_name:Profile
partition_name:p8
table_rows:1085334
***************************10.row***************************
table_name:Profile
partition_name:p9
table_rows:982788
10 rows in set(0.80 sec)
```

[illegible] [illegible]HASH[illegible]ID[illegible]

[illegible]

```
mysql[illegible]EXPLAIN PARTITIONS SELECT*FROM Profile WHERE id=1\G;

***************************1.row***************************

id:1

select_type:SIMPLE

table:Profile

partitions:p1

type:const

possible_keys:PRIMARY

key:PRIMARY

key_len:4

ref:const

rows:1

Extra:

1 row in set(0.00 sec)
```

[illegible]p1[illegible]Profile[illegible]nickname[illegible]EXPLAIN PARTITIONS[illegible]

```
mysql[illegible]EXPLAIN PARTITIONS

-[illegible]SELECT*FROM Profile WHERE nickname='david'\G;

***************************1.row***************************

id:1

select_type:SIMPLE

table:Profile
```

```
partitions:p0,p1,p2,p3,p4,p5,p6,p7,p8,p9

type:ref

possible_keys:nickname

key:nickname

key_len:62

ref:const

rows:10

Extra:Using where

1 row in set(0.00 sec)
```

[illegible]MySQL[illegible]

```
mysql[illegible]SELECT*FROM Profile WHERE nickname='david'\G;

***************************1.row***************************

id:5566

nickname:david

password:3e35d1025659d07ae28e0069ec51ab92

sex:M

rdate:2003-09-20

1 row in set(1.05 sec)
```

[illegible]1.05[illegible]IO[illegible]20[illegible]30[illegible]SQL[illegible]0.26[illegible]

[illegible]InnoDB[illegible]OLTP[illegible]OLTP[illegible]

4.8.6 [illegible]

MySQL 5.6[illegible]ALTER TABLE…EXCHANGE PARTITION[illegible]

[illegible]ALTER TABLE…EXCHANGE PARTITION[illegible]

❑[illegible]

❑[illegible]

❑[illegible]

❑[illegible]ALTER[illegible]INSERT[illegible]CREATE[illegible]DROP[illegible]

[illegible]

❑[illegible]

❑AUTO_INCREMENT[illegible]

[illegible]RANGE[illegible]e[illegible]

```
CREATE TABLE e(
id INT NOT NULL,
fname VARCHAR(30),
lname VARCHAR(30)
)
PARTITION BY RANGE(id)(
PARTITION p0 VALUES LESS THAN(50),
PARTITION p1 VALUES LESS THAN(100),
PARTITION p2 VALUES LESS THAN(150),
PARTITION p3 VALUES LESS THAN(MAXVALUE)
);

INSERT INTO e VALUES

(1669,"Jim","Smith"),

(337,"Mary","Jones"),

(16,"Frank","White"),

(2005,"Linda","Black");
```

[illegible]e2[illegible]e2[illegible]e[illegible]e2[illegible]

```
mysql＞CREATE TABLE e2 LIKE e;

Query OK,0 rows affected(1.34 sec)

mysql＞ALTER TABLE e2 REMOVE PARTITIONING;

Query OK,0 rows affected(0.90 sec)

Records:0 Duplicates:0 Warnings:0
```

[illegible]

```
mysql＞SELECT PARTITION_NAME,TABLE_ROWS

-＞FROM INFORMATION_SCHEMA.PARTITIONS

-＞WHERE TABLE_NAME='e';

+----------------+------------+

|PARTITION_NAME|TABLE_ROWS|

+----------------+------------+

|p0|1|

|p1|0|

|p2|0|

|p3|3|

+----------------+------------+

4 rows in set(0.00 sec)
```

□□□e2□□□□□□□□□□□□□□□e□□□p0□□□□□□□□e2□□

```
mysql□ALTER TABLE e EXCHANGE PARTITION p0 WITH TABLE e2;

Query OK,0 rows affected(0.28 sec)
```

□□□□□□e□□□□□□□□□□□p0□□□□□□□□□□

```
mysql□SELECT PARTITION_NAME,TABLE_ROWS

-□FROM INFORMATION_SCHEMA.PARTITIONS

-□WHERE TABLE_NAME='e';

+----------------+------------+

|PARTITION_NAME|TABLE_ROWS|

+----------------+------------+

|p0|0|

|p1|0|

|p2|0|

|p3|3|

+----------------+------------+

4 rows in set(0.00 sec)
```

□□□□□□□e2□□□□□□□□□□□

```
mysql□SELECT*FROM e2;

+----+-------+-------+

|id|fname|lname|

+----+-------+-------+

|16|Frank|White|

+----+-------+-------+

1 row in set(0.00 sec)
```

4.9 [illegible]

[illegible]InnoDB[illegible]InnoDB[illegible]MySQL[illegible]InnoDB[illegible]MySQL[illegible]

[illegible]MySQL[illegible]RANGE[illegible]LIST[illegible]HASH[illegible]KEY[illegible]COLUMNS[illegible]HASH[illegible]KEY[illegible]OLTP[illegible]

□5□ □□□□□

□□□

□□□□□□□□□□□□□□□□□□——□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□DBA□□□□□DBA□□□□□□□□□□□□□□□□□□□□□□□□□□□SQL□□MySQL□□□iostat□□□□□□□□□□□100%□□20%□□□□□□□□□□□□□□□□□□□

□□□□□□□□InnoDB□□MySQL□□□□□□□□□□□□□□□□□□□□□□□□□□□

5.1 InnoDB[illegible]

InnoDB[illegible]

❑B+[illegible]

❑[illegible]

❑[illegible]

[illegible]InnoDB[illegible]InnoDB[illegible]

B+[illegible]B+[illegible]Key Value[illegible]

[illegible] B+[illegible]B[illegible]binary[illegible]balance[illegible]B+[illegible]B+[illegible]

[illegible]DBA[illegible]B+[illegible]B+[illegible]

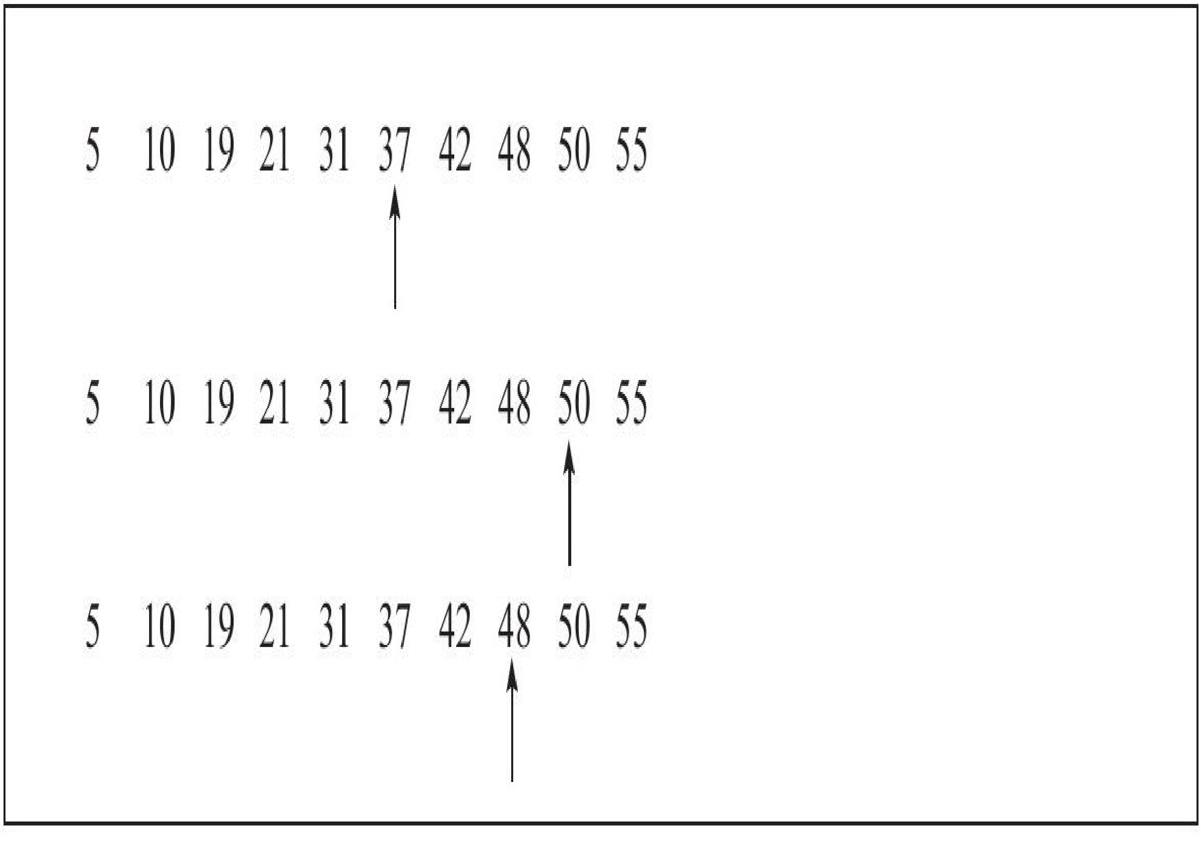
5.2 [illegible]

B+[illegible]B+[illegible]

5.2.1 [illegible]

[illegible]binary search[illegible]

[illegible]5[illegible]10[illegible]19[illegible]21[illegible]31[illegible]37[illegible]42[illegible]48[illegible]50[illegible]52[illegible]10[illegible]10[illegible]48[illegible]5-1[illegible]



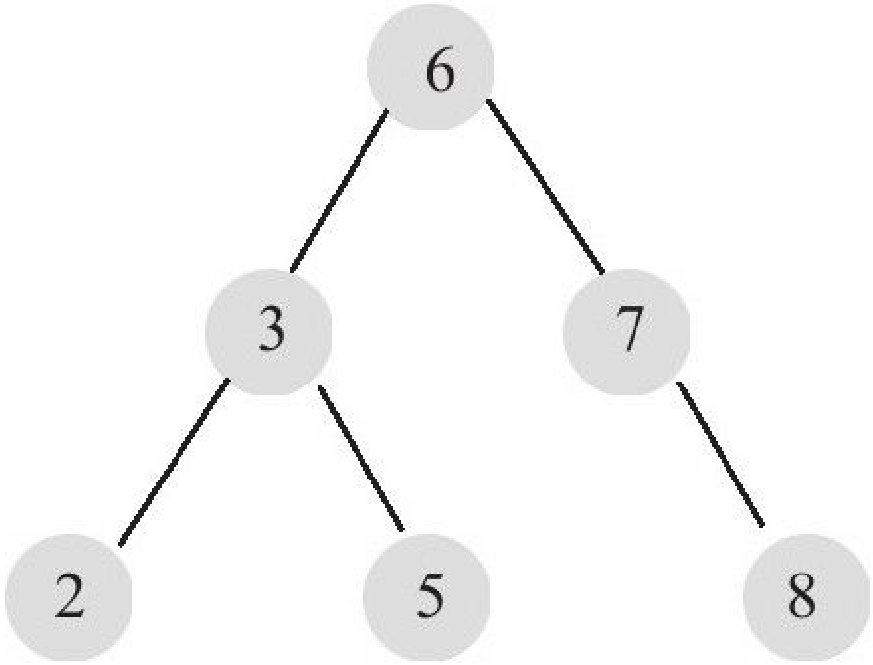
[illegible] 5-1 [illegible]

[illegible]5-1[illegible]3[illegible]48[illegible]8[illegible]5[illegible]1[illegible]4[illegible]10[illegible]1+2+3+4+5+6+7+8+9+10[illegible]/10=5.5[illegible]4+3+2+4+3+1+4+3+2+3[illegible]/10=2.9[illegible]10[illegible]4[illegible]

[illegible]1946[illegible]1962[illegible]Page Directory[illegible]Page Directory[illegible]

5.2.2 □□□□□□□□□□□

□□□B+□□□□□□□□□□□□□□□□B+□□□□□□□□□□□□□□□□□□B□□5-2□□□□□□□□□□□



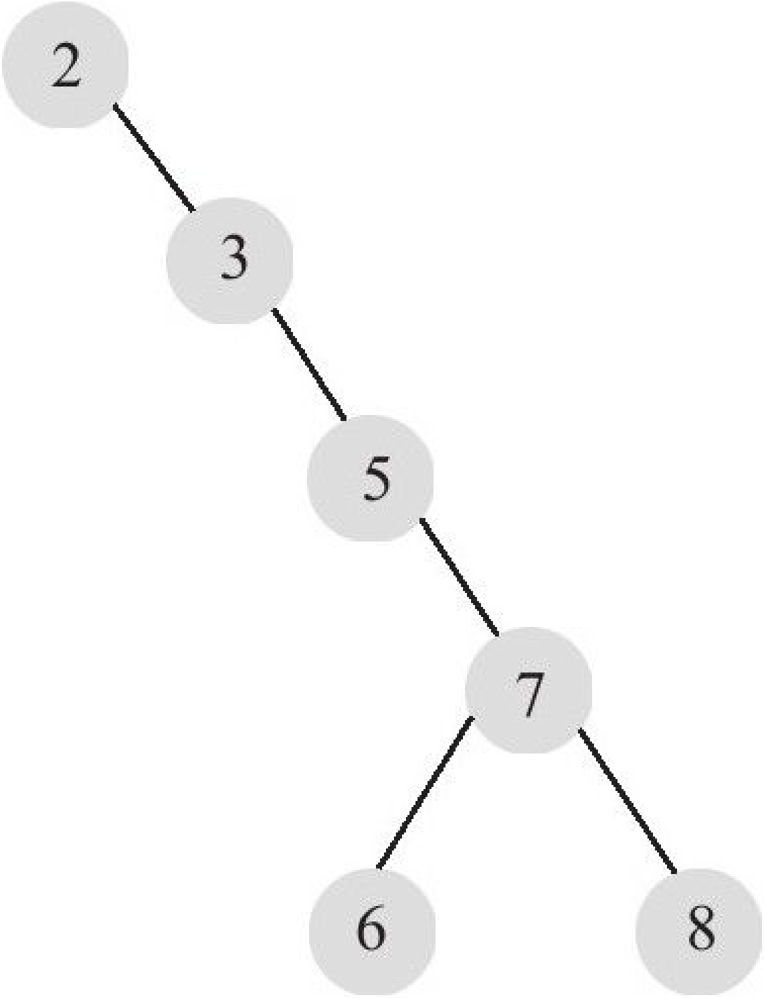
□ 5-2 □□□□□

□5-2□□5-2□□□□□□□□□□□□□□□□2□3□5□6□7□8□

□□5-2□□□□□□□□□□□□□□□□5□□□□□□□□□□□□□6□6□□5□□□□□6□□□□□□□3□□5□□3□□□□□□□□□□□□3□□□□□2□3□5□6□7□8□□□□□□□□□3□□□□□□□□□□□□□□8□□□□□□□□□□3□□□□□□□□□□□6□□□□□□□□□□□□□□□□□□□□□□□□□□1+2+3+4+5+6□/6=3.3□□□□□□□□□□□□□□□□3+3+3+2+2+1□/6=2.3□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□2□3□5□6□7□8□□□□□□□□□□□□5-3□□□□□□□□□□□

[illegible]5-3[illegible]1+2+3+4+5+5[illegible]/6=3.16[illegible]——[illegible]AVL[illegible]

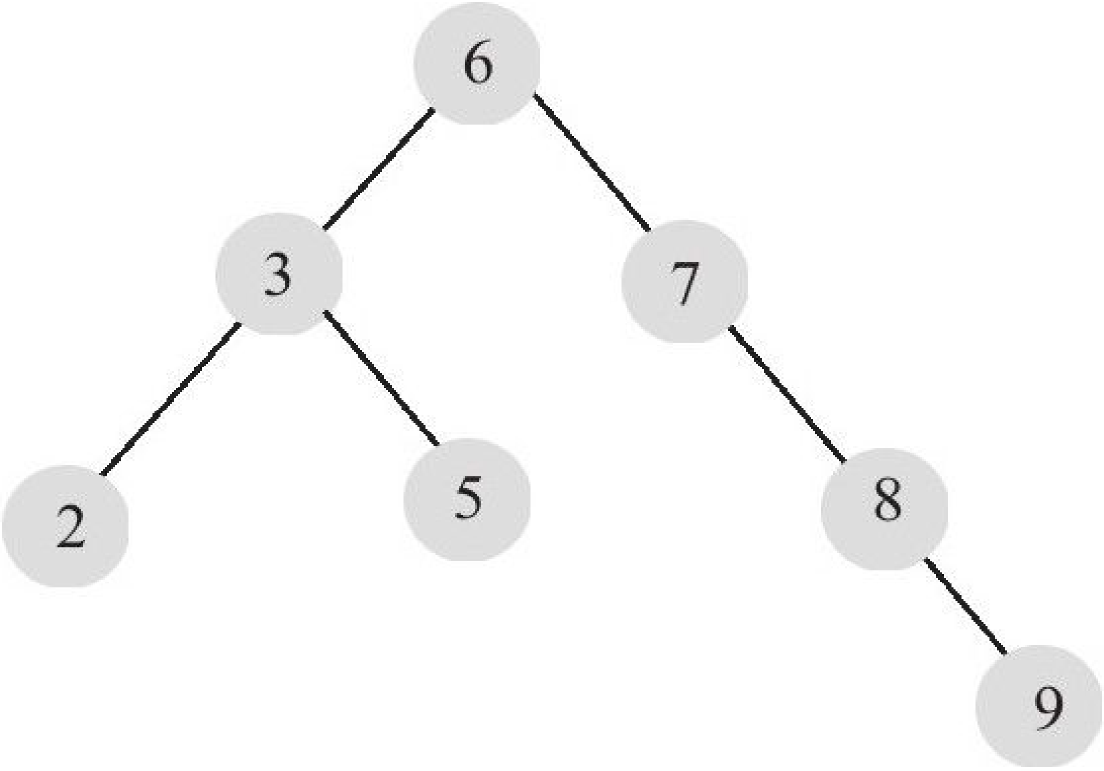


[illegible] 5-3 [illegible]

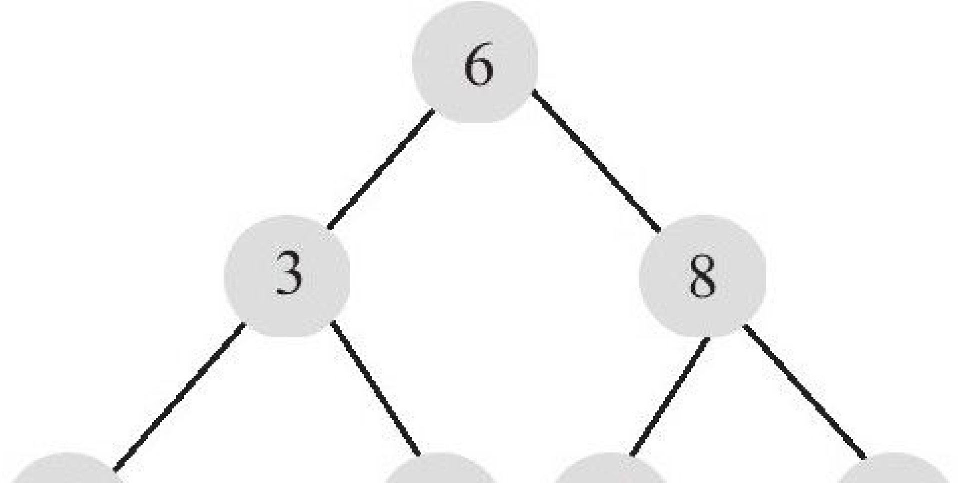
[illegible]1[illegible]5-3[illegible]5-2[illegible]

[illegible]1[illegible]5-2[illegible]9[illegible]

□□5-4□□□□□□



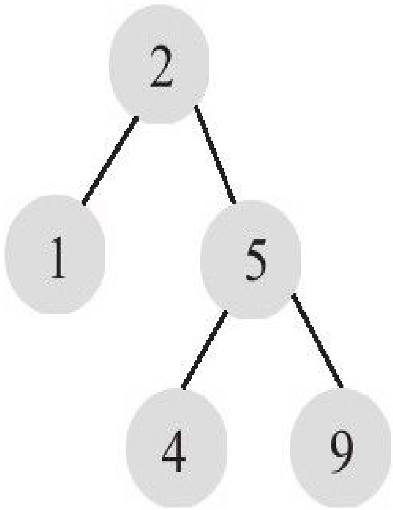
插入新值9



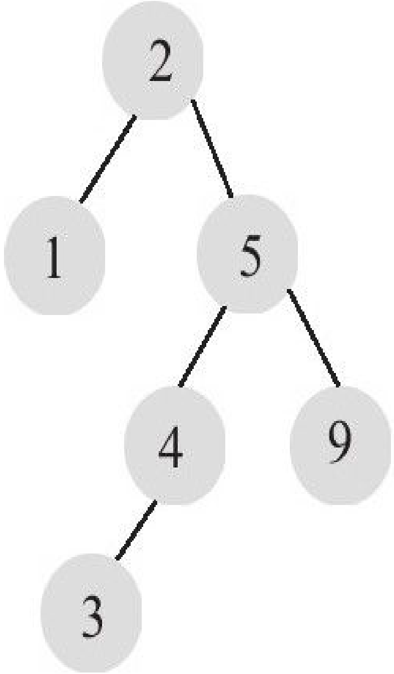
左旋以保证平衡

[illegible] 5-4 [illegible]9[illegible]

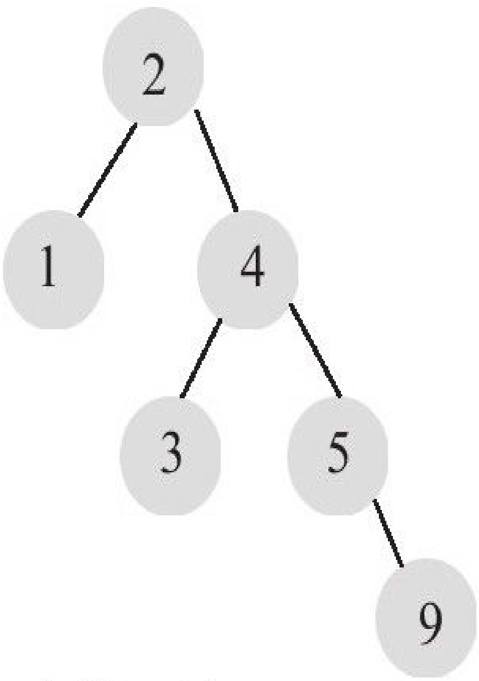
[illegible]5-5[illegible]



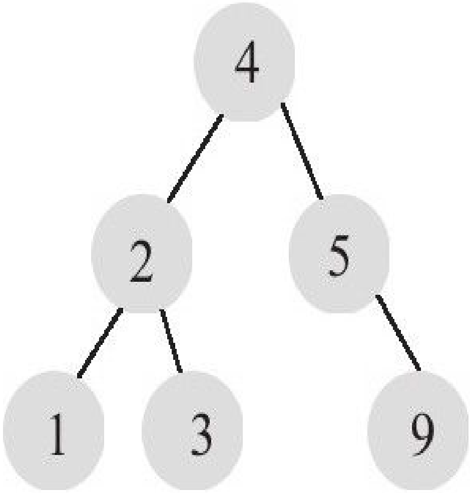
平衡二叉树



插入新键值3



右旋一次



再左旋一次

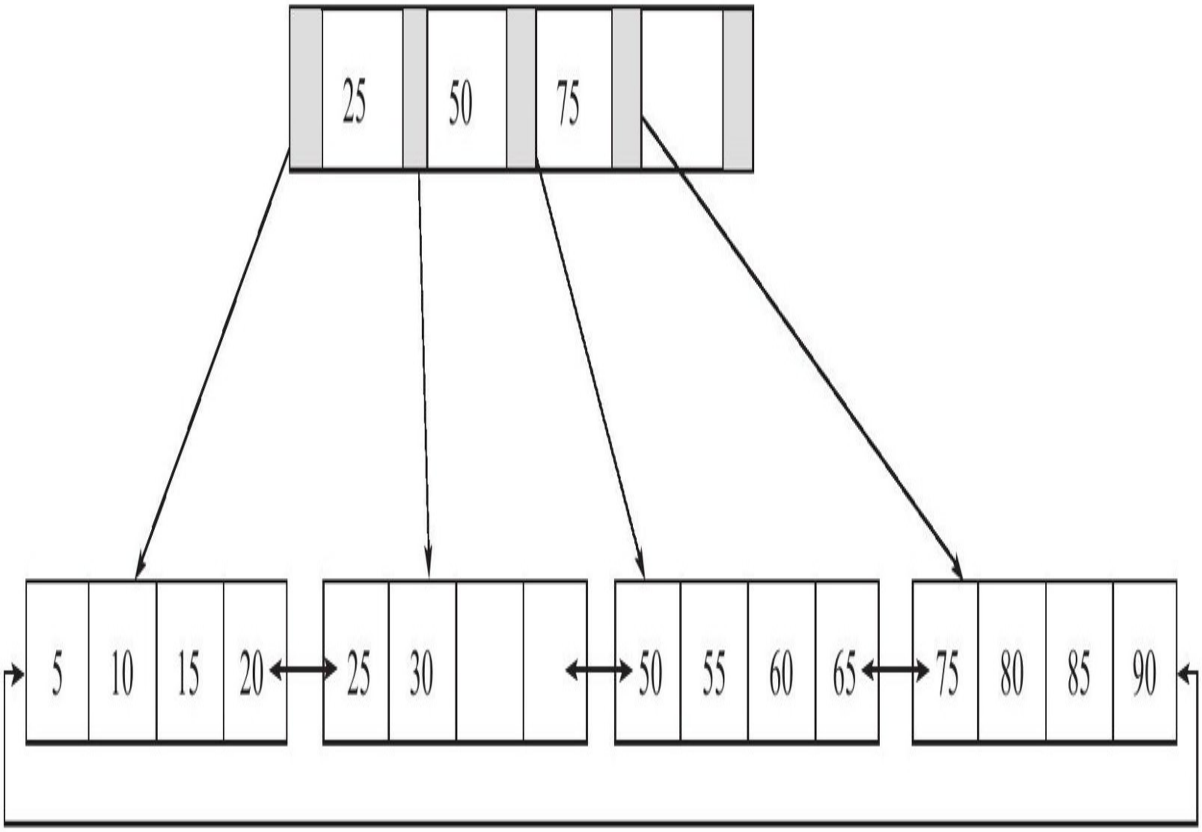
[illegible] 5-5 [illegible]

[illegible]5-4[illegible]5-5[illegible]

5.3 B+□

B+□□□□□□□□□□□□□□□□□□□□□□□□B+□□B□□□□□□□□□□□ISAM□□□□□□□□□□□□□MyISAM□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□B□□□□□□

B+□□□B+□□□□□□B+□□□□□□□□□□□□□□□□□□□□□□□□□□□□B+□□□B+□□□□□□2□□□□□□4□□□□□□□fan out□□5□□□5-6□□□



□ 5-6 □□□□□2□B+□

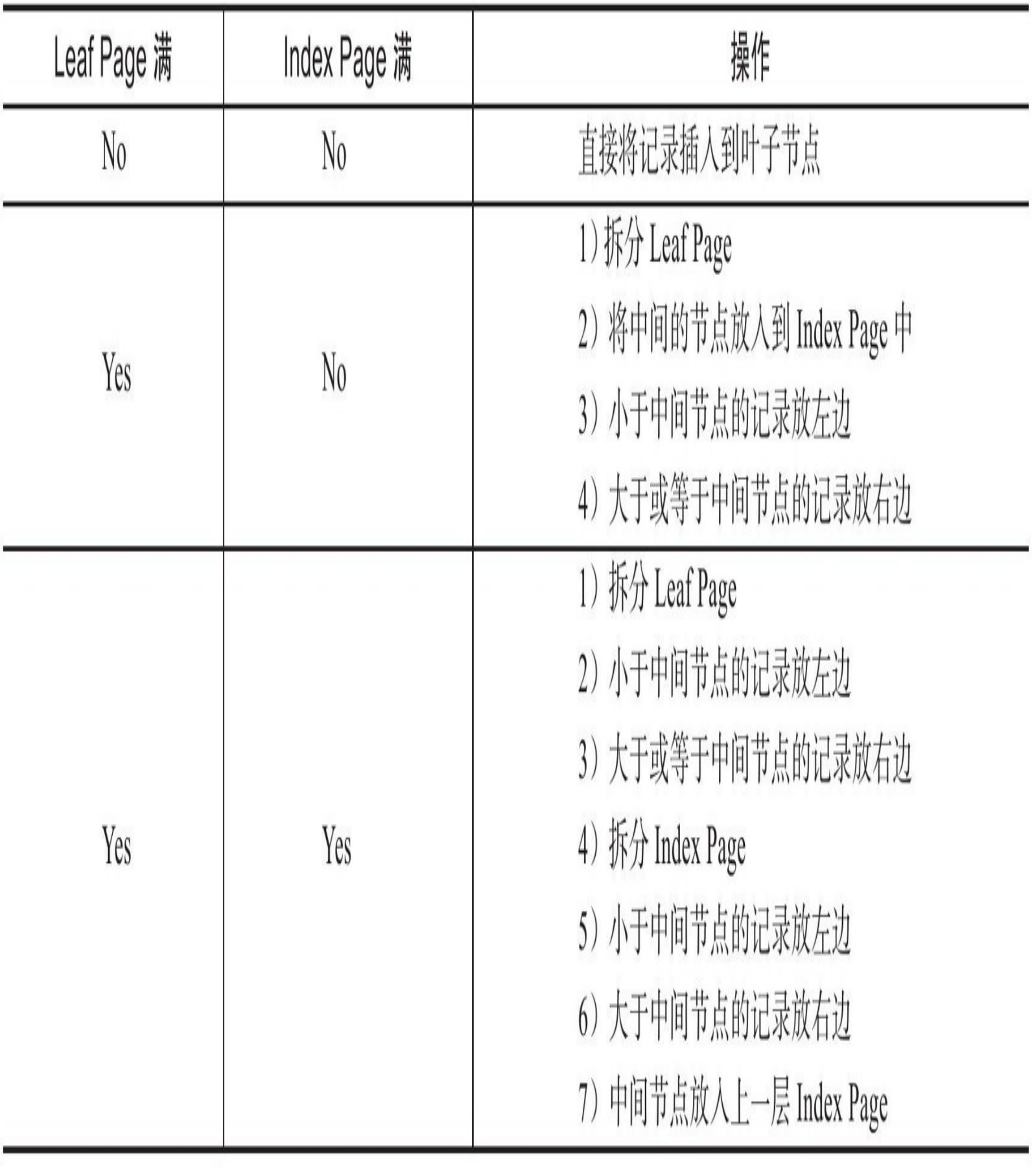
□□5-6□□□5□10□15□20□25□30□50□55□60□65□75□80□

85[illegible]90[illegible]

5.3.1 B+[illegible]

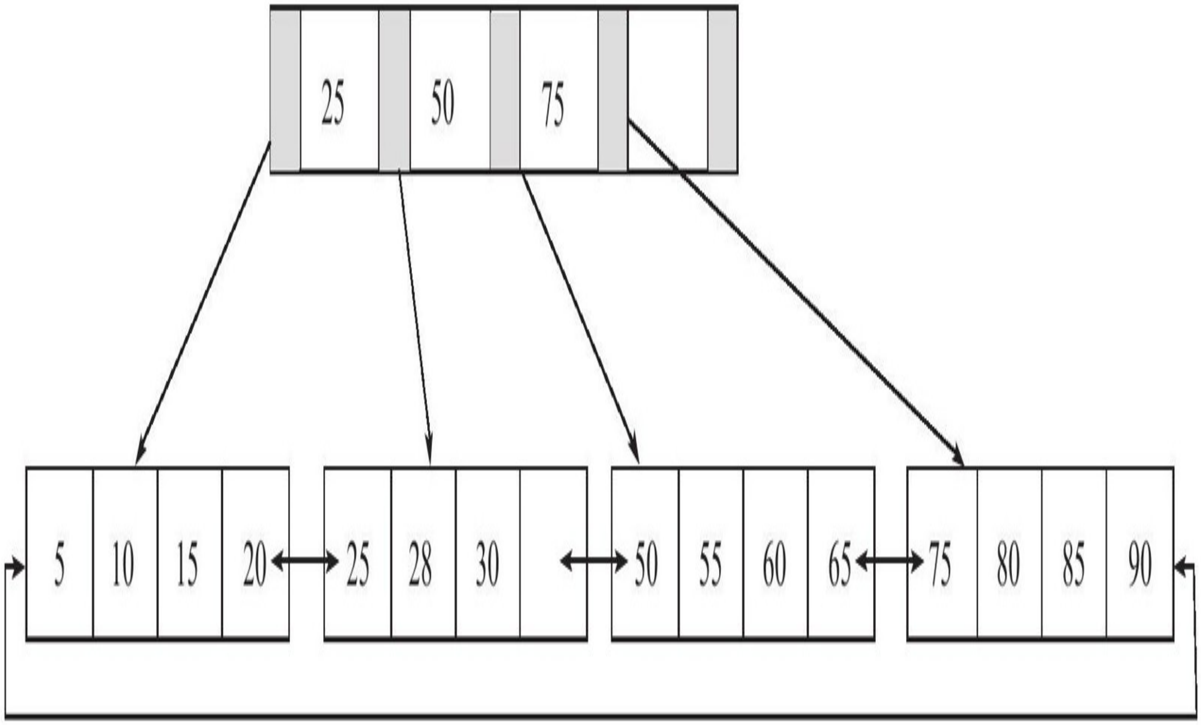
B+[illegible]B+[illegible]5-1[illegible]

表 5-1 B+ 树插入的 3 种情况



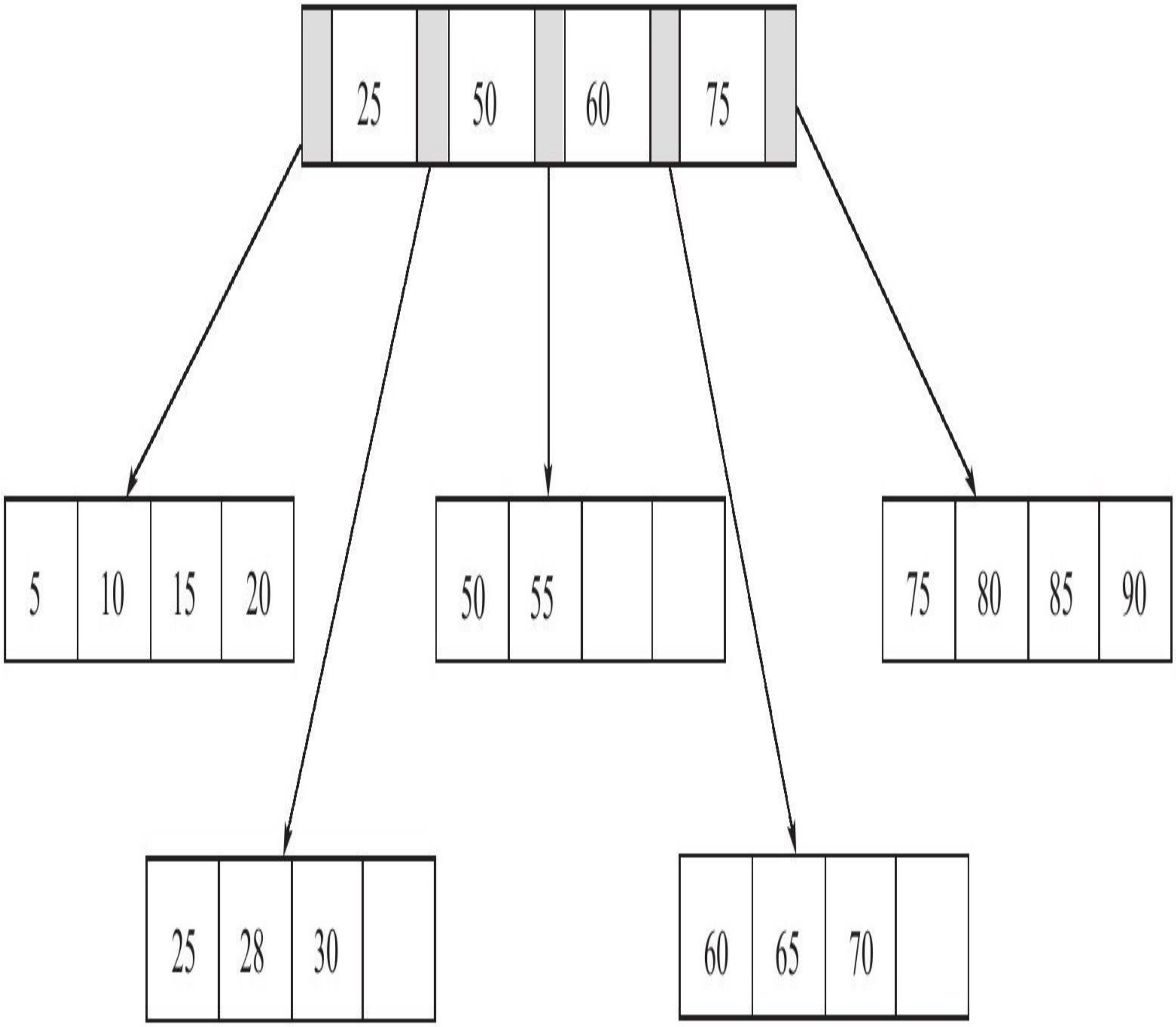
Leaf Page 满	Index Page 满	操作
No	No	直接将记录插入到叶子节点
Yes	No	1）拆分 Leaf Page 2）将中间的节点放入到 Index Page 中 3）小于中间节点的记录放左边 4）大于或等于中间节点的记录放右边
Yes	Yes	1）拆分 Leaf Page 2）小于中间节点的记录放左边 3）大于或等于中间节点的记录放右边 4）拆分 Index Page 5）小于中间节点的记录放左边 6）大于中间节点的记录放右边 7）中间节点放入上一层 Index Page

[illegible]B+[illegible]5-6[illegible]B+[illegible]28[illegible]Leaf Page[illegible]Index Page[illegible]5-7[illegible]



[illegible] 5-7 [illegible]28

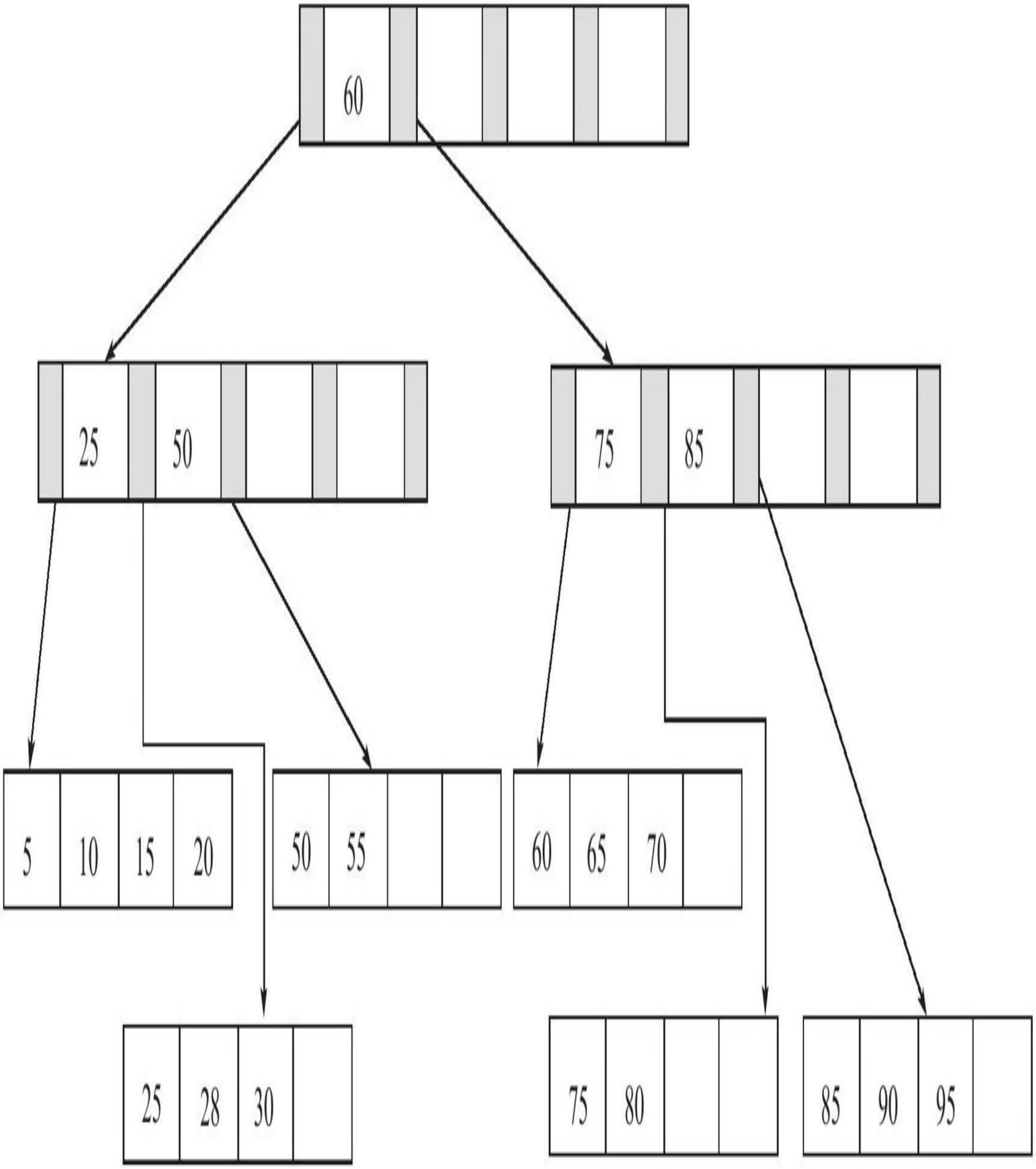
[illegible]70[illegible]Leaf Page[illegible]Index Page[illegible]5-1[illegible]Leaf Page[illegible]55[illegible]55[illegible]60[illegible]65[illegible]70[illegible]60[illegible]5-8[illegible]



[illegible] 5-8 [illegible]70

[illegible]5-6[illegible]5-7[illegible]

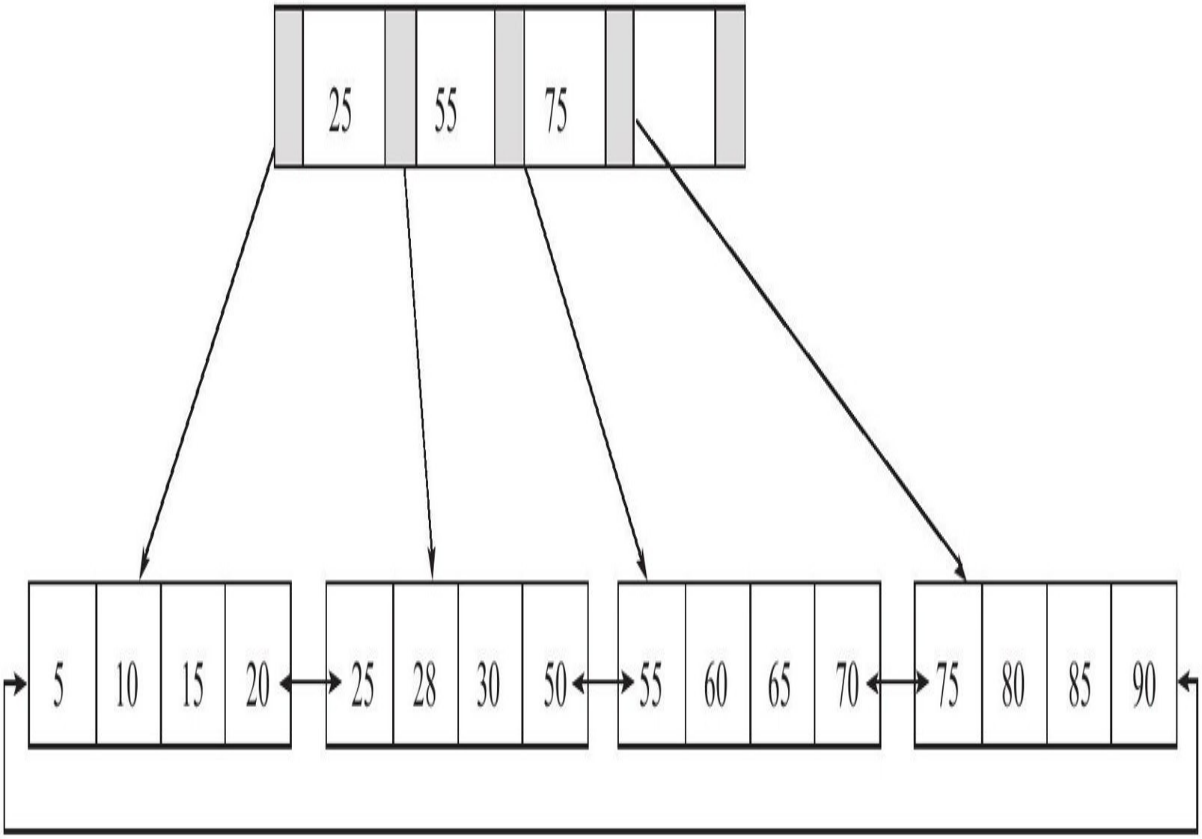
[illegible]95[illegible]5-1[illegible]Leaf Page[illegible]Index Page[illegible]5-9[illegible]



□ 5-9 □□□□95

□□□□□□□□□□□□B+□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□split□□□□□□B+□□B+□□□□□□□□□□□□□□□□□□Rotation□□□□

□□□□□Leaf Page□□□□□□□□□□□□□□□□□□□□□□□□B+□□□□□□□□□□□ □□□ 5-7□□□□□□□□□70□□□B+□□□□□□□□□□□□□□□□□□□□□□□□□□□5-10□ □□□□□



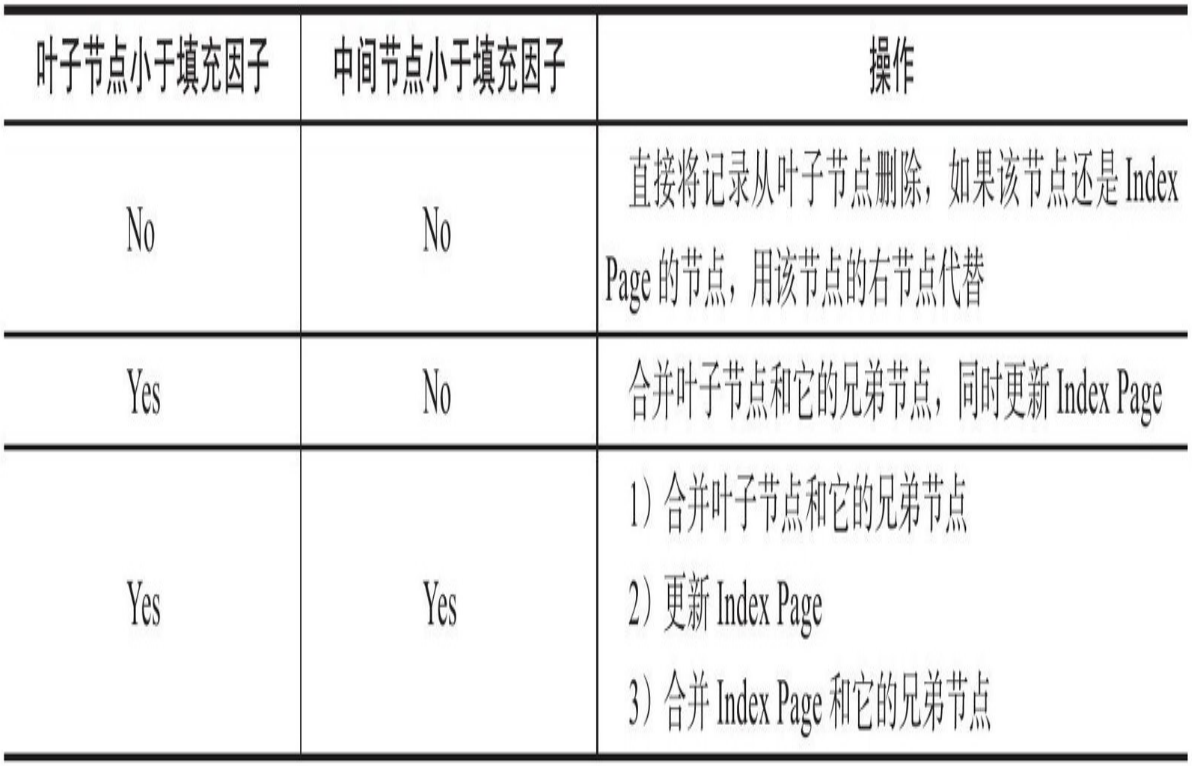
□ 5-10 B+□□□□□□

□□5-10□□□□□□□□□□□□B+□□□□□□□□□□□□□□□□□B+□□□□□□□□2□

5.3.2 B+[illegible]

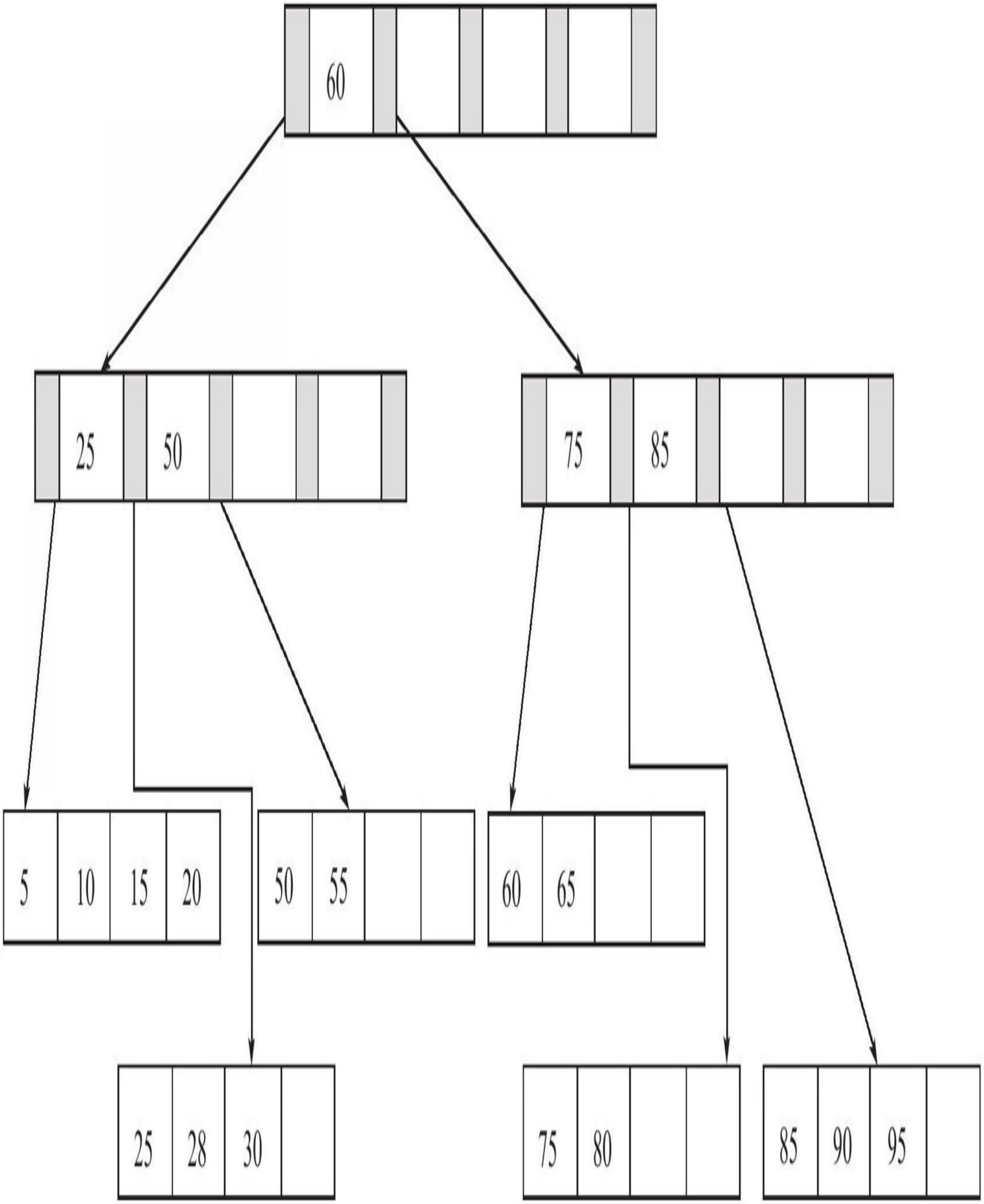
B+[illegible]fill factor[illegible]50%[illegible]B+[illegible]B+[illegible]5-2[illegible]

表 5-2 B+ 树删除操作的三种情况



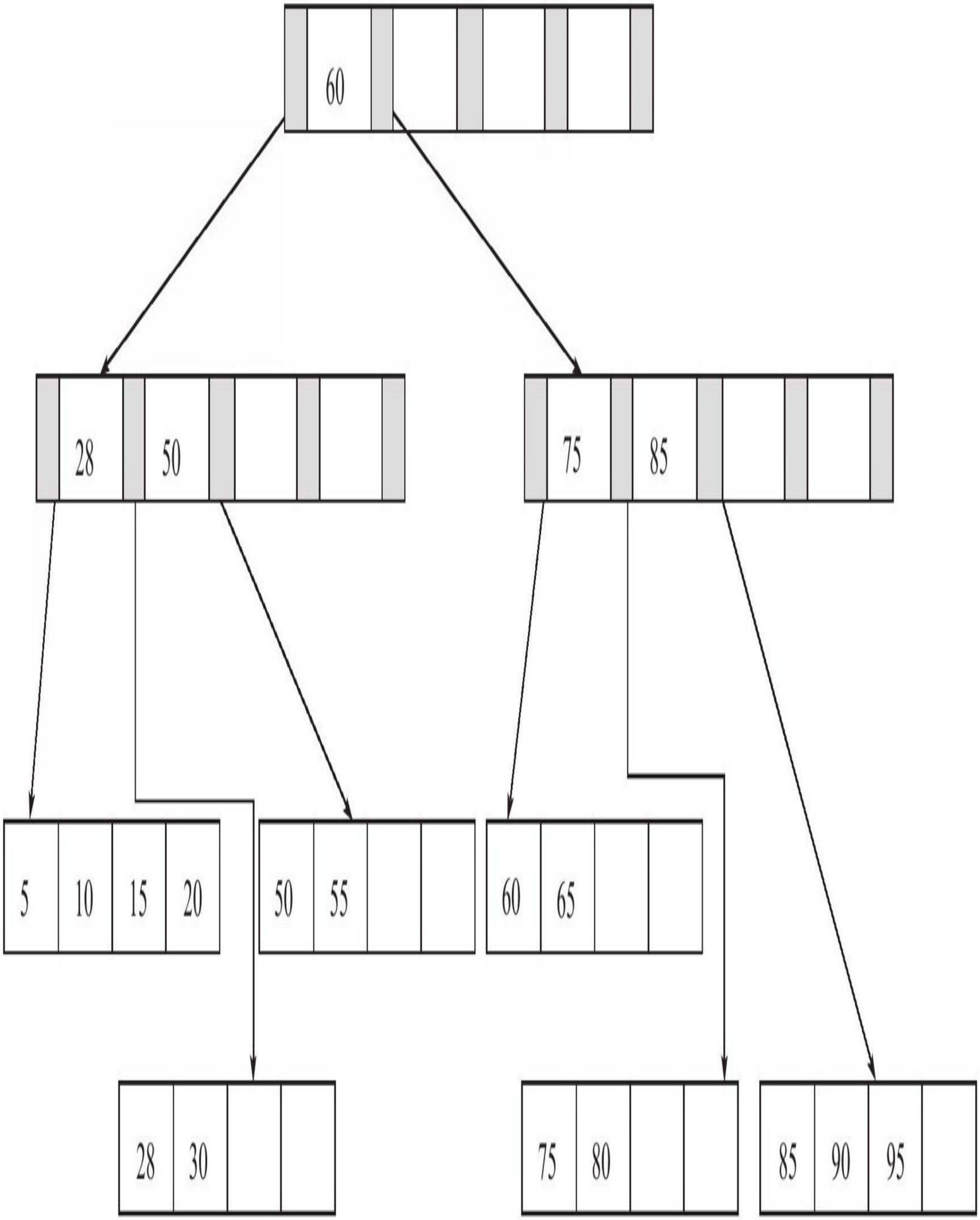
叶子节点小于填充因子	中间节点小于填充因子	操作
No	No	直接将记录从叶子节点删除，如果该节点还是 Index Page 的节点，用该节点的右节点代替
Yes	No	合并叶子节点和它的兄弟节点，同时更新 Index Page
Yes	Yes	1）合并叶子节点和它的兄弟节点 2）更新 Index Page 3）合并 Index Page 和它的兄弟节点

[illegible]5-9[illegible]B+[illegible]70[illegible]5-2[illegible]5-11[illegible]



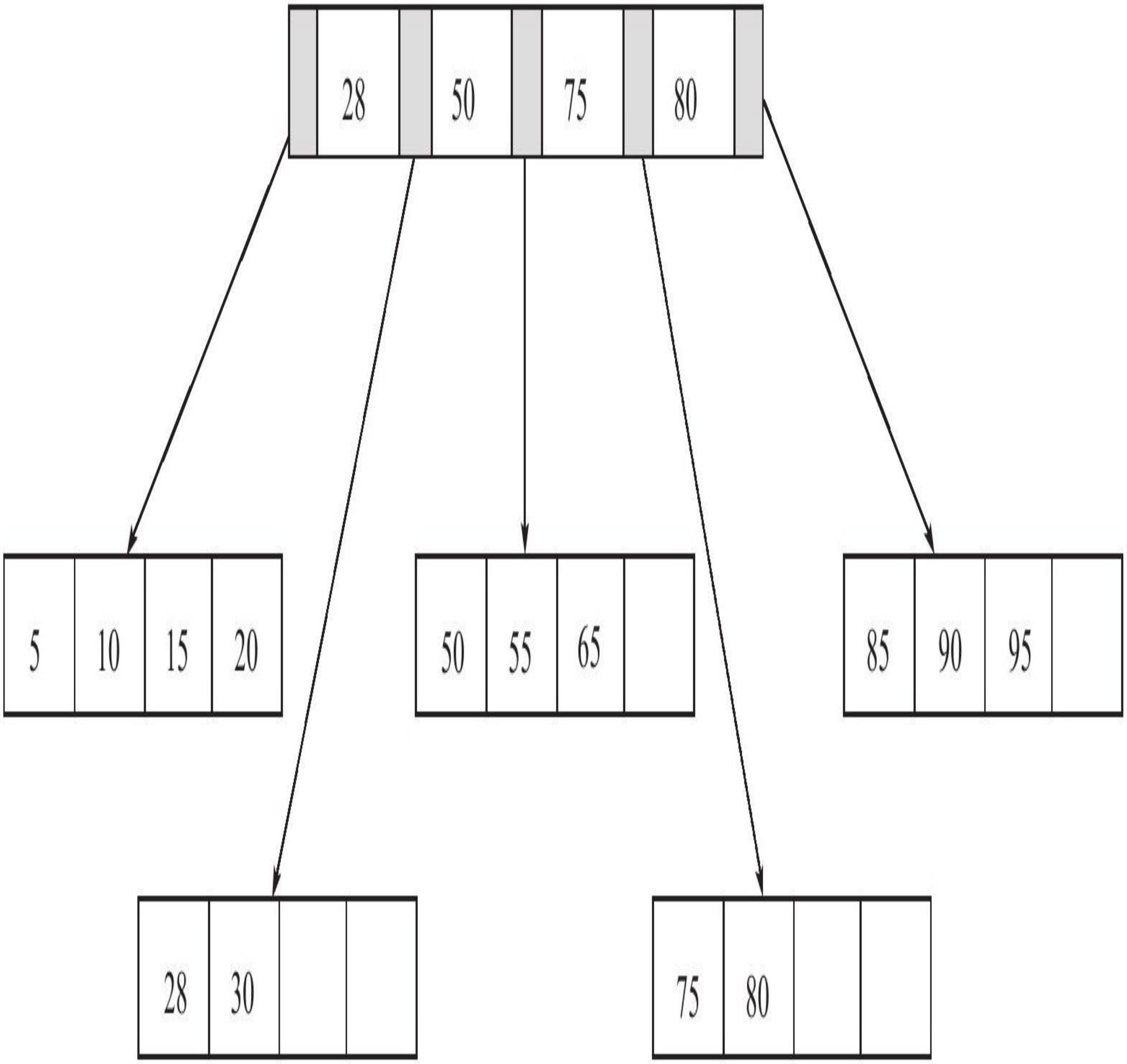
[illegible] 5-11 [illegible]70

□□□□□□□25□□□□□□□□5-2□□□□□□□□□□□□□□□IndexPage□□□□□□□□
□Leaf Page□□25□□□□□25□□□□□□□28□□□Page Index□□□□□□□5-
12□



[illegible] 5-12 [illegible]25

□□□□□□□□60□□□□□□Leaf Page□□□□60□□□□□Fill Factor□□50%□□□□□□□□□□□□□□□□□Index Page□□□□□□□□□Index Page□□□□□□□□□□□5-13□



□ 5-13 □□□□60

5.4 B+□□□

□□□□□□□B+□□□□□□□□□□□□□B+□□□□□□□□B+□□□□□□□□□□□□B+□□□□□□□□□□□□□□□□□□□□□□□□□□B+□□□□□□□□2□4□□□□□□□□□□□□□□□□□□□□□□□2□4□IO□□□□□□□□□□□□□□□□□□□□□□□□100□IO□2□4□□IO□□□□□□□□□0.02□0.04□□

□□□□□B+□□□□□□□□□□□□clustered inex□□□□□□□secondary index□[1]□□□□□□□□□□□□□□□□□□□□□B+□□□

5.4.1 □□□□

□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□clustered index□□□□□□□□□□□□□□□B+□□B+□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□B+□□□B+□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
CREATE TABLE t□

a INT NOT NULL,

b VARCHAR(8000),

c INT NOT NULL,

PRIMARY KEY(a),

KEY idx_c(c)

)ENGINE=INNODB;

INSERT INTO t SELECT 1,REPEAT('a',7000),-1;

INSERT INTO t SELECT 2,REPEAT('a',7000),-2;
INSERT INTO t SELECT 3,REPEAT('a',7000),-3;
INSERT INTO t SELECT 4,REPEAT('a',7000),-4;
```

□□□□□□□□□□□b□□□7000□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□py_innodb_page_info□□□□□□□□□□□□

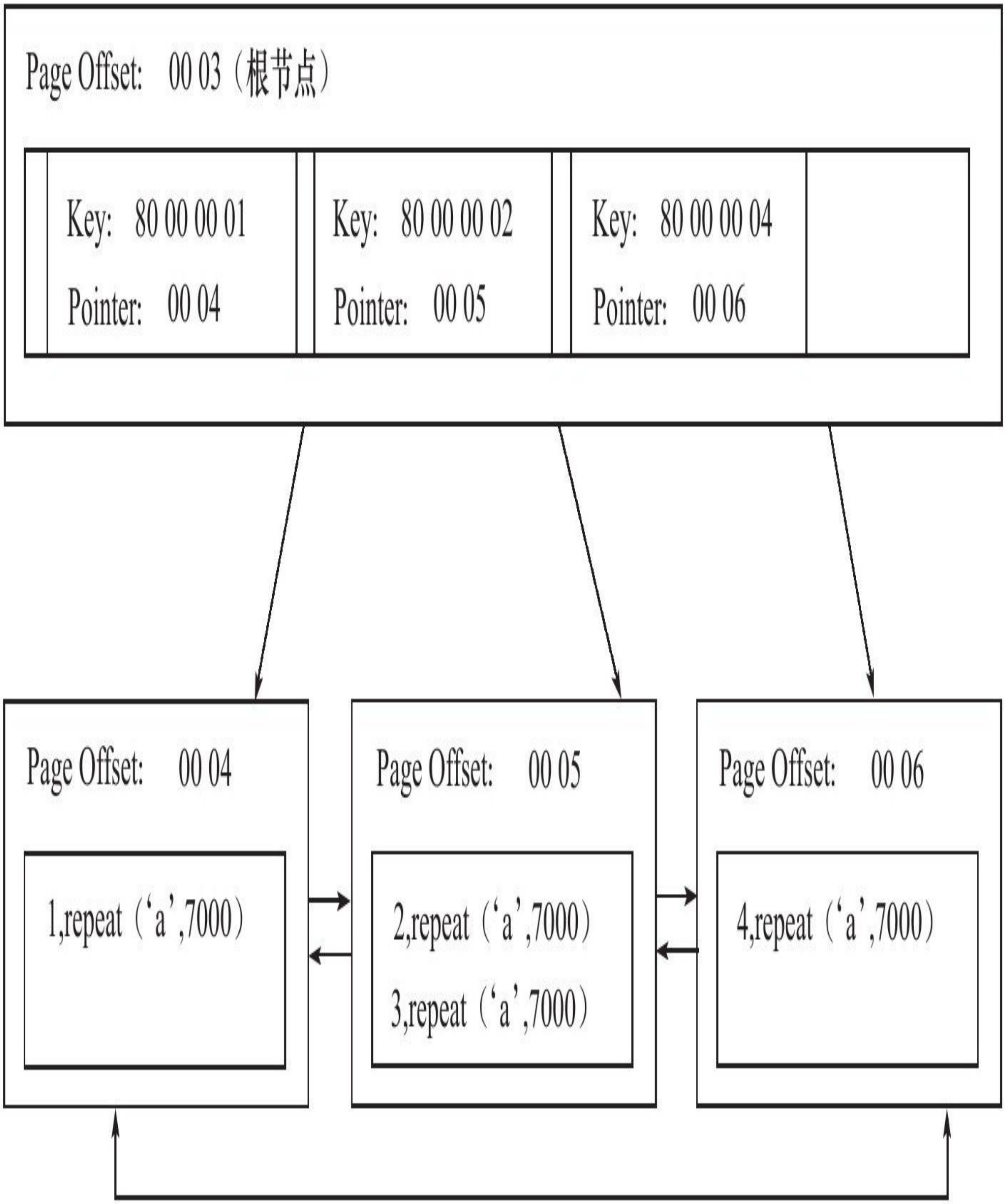
```
[root@nineyou0-43 data]#py_innodb_page_info.py-v mytest/t.ibd
page offset 00000000,page type□File Space Header□
page offset 00000001,page type□Insert Buffer Bitmap□
page offset 00000002,page type□File Segment inode□
page offset 00000003,page type□B-tree Node□,page level□0001□
page offset 00000004,page type□B-tree Node□,page level□0000□
page offset 00000005,page type□B-tree Node□,page level□0000□
page offset 00000006,page type□B-tree Node□,page level□0000□
page offset 00000000,page type□Freshly Allocated Page□
Total number of page:8:
Freshly Allocated Page:1
Insert Buffer Bitmap:1
File Space Header:1
B-tree Node:4
File Segment inode:1
```

page level□0000□□□page level□0001□□□□□□□□□□B+□□□□2□□□□□B+□□□□□□hexdump□□□□□□□□□□□□□□□□□□

```
0000c000 c2 33 62 95 00 00 00 03 ff ff ff ff ff ff ff ff|.3b............|
0000c010 00 00 00 0a b6 8c ce 57 45 bf 00 00 00 00 00 00|.......WE.......|
0000c020 00 00 00 00 00 f9 00 02 00 a2 80 05 00 00 00 00|...............|
0000c030 00 9a 00 02 00 02 00 03 00 00 00 00 00 00 00 00|...............|
0000c040 00 01 00 00 00 00 00 00 01 e2 00 00 00 f9 00 00|................|
0000c050 00 02 00 f2 00 00 00 f9 00 00 00 02 00 32 01 00|............2..|
0000c060 02 00 1b 69 6e 66 69 6d 75 6d 00 04 00 0b 00 00|...infimum......|
0000c070 73 75 70 72 65 6d 75 6d 00 10 00 11 00 0e 80 00|supremum........|
0000c080 00 01 00 00 00 04 00 00 00 19 00 0e 80 00 00 02|................|
0000c090 00 00 00 05 00 00 00 21 ff d6 80 00 00 04 00 00|.......!........|
0000c0a0 00 06 00 00 00 00 00 00 00 00 00 00 00 00 00 00|................|
0000c0b0 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00|................|
0000c0c0 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00|................|
......
0000fff0 00 00 00 00 00 70 00 63 73 d8 52 3a b6 8c ce 57|.....p.cs.R:...W|
```

□□□□□□□□□□□Page Directory□□□□□□□00 63□□□□□□□□□□□□□□□□□□Recorder Header□□□□0xc063□□□□□69 6e 66 69 6d 75 6d 00□□□□infimum□□□□□□□□5□□01 00 02 00 1b□□Recorder Header□□□□□4□□8□□□□1□□□□□□□□□□□□□□□□□□□□InnoDB□Page Directory□□□□□□□infimum□□□□□□□Recorder Header□□□□□□□□00 1b□□□□□□□□□□□□□c063+1b=c07e□□□□□□□80 00 00 01□□□□□□□1□□□□□□□□INT□□□□□□□□□□□□0x80 00 00 01□□□□0x0001□□80 00 00 01□□□00 00 00 04□□□□□□□□□□□□□□□□□□□□80 00 00 02□80 00 00 04□□□□□□□□□□□□□□□□

□□□5-14□□□


Page Offset: 00 03（根节点）
Key: 80 00 00 01
Pointer: 00 04
Key: 80 00 00 02
Pointer: 00 05
Key: 80 00 00 04
Pointer: 00 06
Page Offset: 00 04
1,repeat（‘a’,7000）
Page Offset: 00 05
2,repeat（‘a’,7000）
3,repeat（‘a’,7000）
Page Offset: 00 06
4,repeat（‘a’,7000）

[illegible] 5-14 B+[illegible]

[illegible]5-14[illegible]

[illegible]

[illegible]10[illegible]B+[illegible]10[illegible]EXPLAIN[illegible]

```
mysql[illegible]EXPLAIN

-[illegible]SELECT*FROM Profile ORDER BY id LIMIT 10\G;

***************************1.row***************************

id:1

select_type:SIMPLE

table:Profile

type:index

possible_keys:NULL

key:PRIMARY

key_len:4

ref:NULL

rows:10

Extra:

1 row in set(0.00 sec)
```

[illegible]ORDER BY[illegible]filesort[illegible]

[illegible]range query[illegible]

```
mysql[illegible]EXPLAIN

-[illegible]SELECT*FROM Profile

-[illegible]WHERE id[illegible]10 AND id[illegible]10000\G;

***************************1.row***************************
```

```
id:1

select_type:SIMPLE

table:Profile

type:range

possible_keys:PRIMARY

key:PRIMARY

key_len:4

ref:NULL

rows:14868

Extra:Using where

1 row in set□0.01 sec□
```

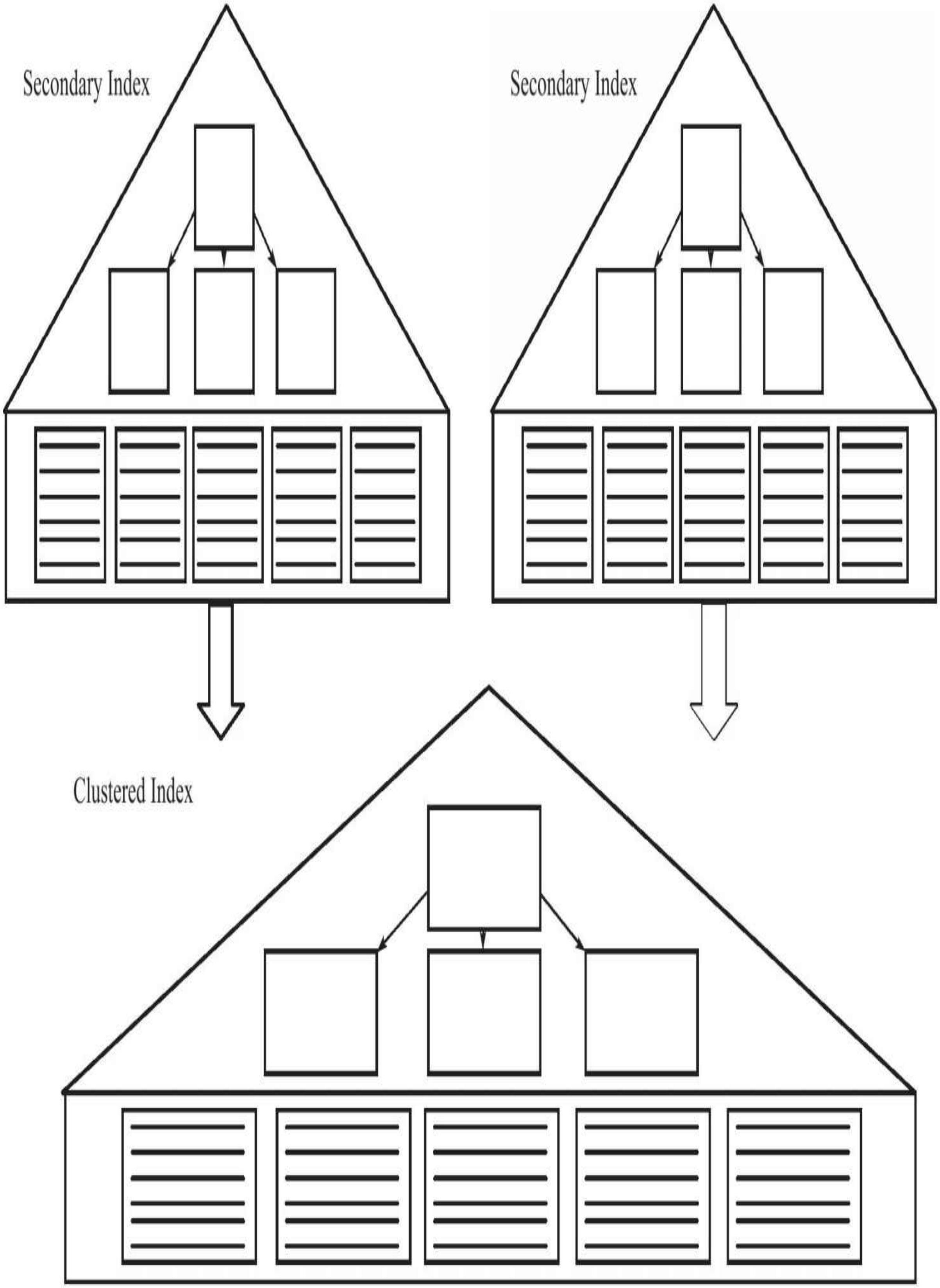
□□EXPLAIN□□□MySQL□□□□□□□□□execute plan□□□□□rows□□□□□□□□□□□□□□□□□□□□□□□□□rows□□□□□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□9946□□□□

```
mysql□SELECT COUNT(*)from Profile

-□WHERE id□10 AND id□10000;

***************************1.row***************************

COUNT(1):9946

1 row in set(0.00 sec)
```

[1] □□□□□□□□□□□□□□non-clustered index□□

5.4.2 [illegible]

[illegible]Secondary Index[illegible]bookmark[illegible]InnoDB[illegible]InnoDB[illegible]InnoDB[illegible]5-15[illegible]InnoDB[illegible]


Secondary Index
Secondary Index
Clustered Index

[illegible] 5-15 [illegible]

[illegible]InnoDB[illegible]3[illegible]3[illegible]3[illegible]3[illegible]6[illegible]IO[illegible]

[illegible]Microsoft SQL Server[illegible]MySQL[illegible]MyISAM[illegible]NOT NULL[illegible]Row Identifiedr[illegible]RID[illegible]“[illegible]”[illegible]

[illegible]Microsoft SQL Server[illegible]DBA[illegible]Microsoft SQL Server[illegible]

[illegible]OLAP[illegible]On-Line Analytical Processing[illegible]B+[illegible]read ahead[illegible]InnoDB[illegible]MyISAM[illegible]It all depends[illegible]

[illegible]InnoDB[illegible]t[illegible]t[illegible]c[illegible]c[illegible]

```
mysql[illegible]ALTER TABLE t

-[illegible]ADD c INT NOT NULL;

Query OK,4 rows affected(0.24 sec)

Records:4 Duplicates:0 Warnings:0

mysql[illegible]UPDATE t SET c=0-a;

Query OK,4 rows affected(0.04 sec)
```

```
Rows matched:4 Changed:4 Warnings:0
mysql>ALTER TALBE t ADDKEY idx_c(c);
Query OK,4 rows affected(0.28 sec)
Records:4 Duplicates:0 Warnings:0
mysql>SHOW INDEX FROM t\G;
***************************1.row***************************
Table:t
Non_unique:0
Key_name:PRIMARY
Seq_in_index:1
Column_name:a
Collation:A
Cardinality:2
Sub_part:NULL
Packed:NULL
Null:
Index_type:BTREE
Comment:
***************************2.row***************************
Table:t
Non_unique:1
Key_name:idx_c
Seq_in_index:1
Column_name:c
Collation:A
Cardinality:2
Sub_part:NULL
Packed:NULL
Null:
Index_type:BTREE
```

```
Comment:

2 rows in set□0.00 sec□

mysql□select a,c from t;

+---+----+

|a|c|

+---+----+

|4|-4|

|3|-3|

|2|-2|

|1|-1|

+---+----+

4 rows in set□0.00 sec□
```

□□□py_innodb_page_info□□□□□□□□□□□□

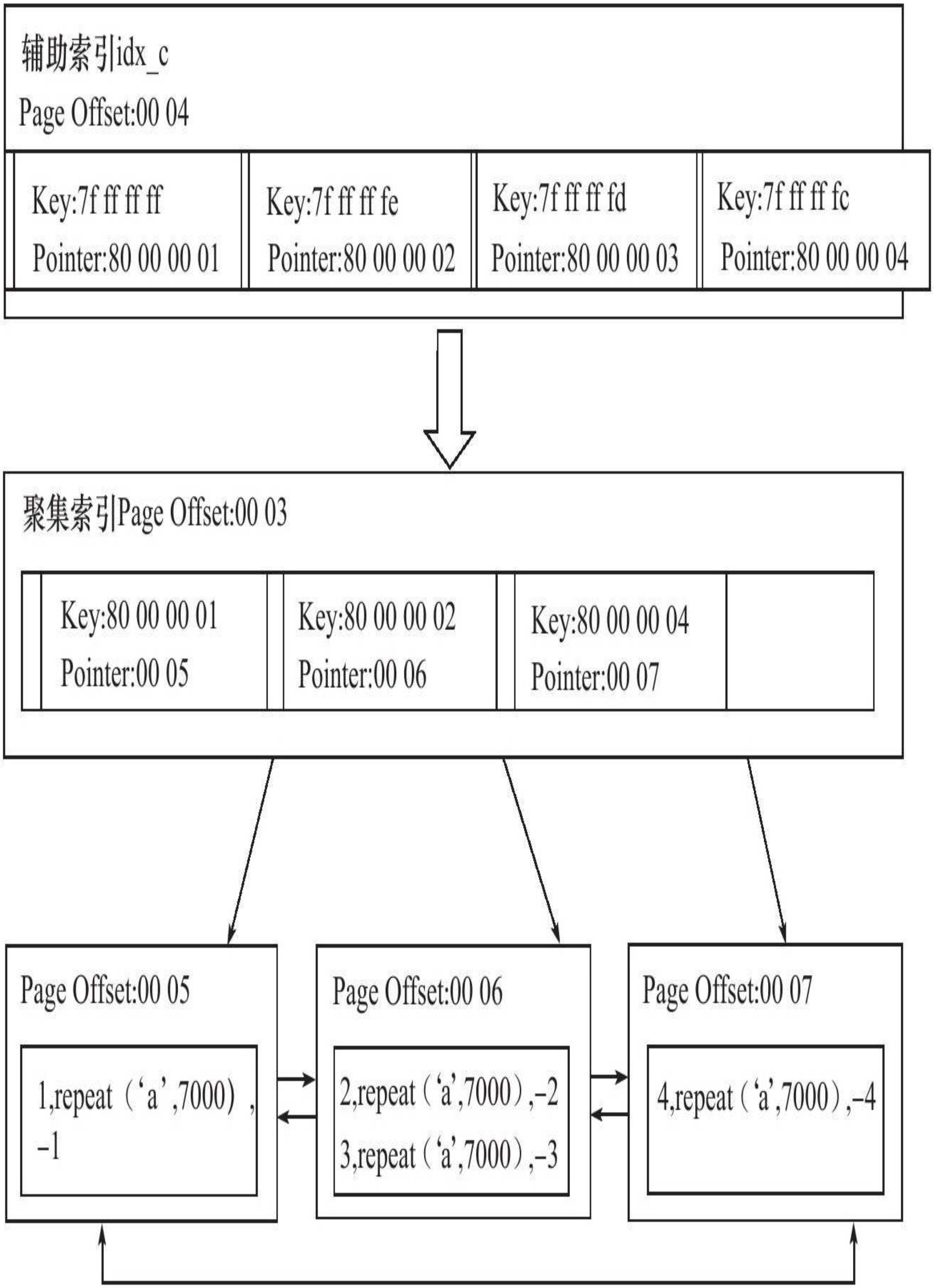
```
[root@nineyou0-43 mytest]#py_innodb_page_info.py-v t.ibd

page offset 00000000,page type□File Space Header□

page offset 00000001,page type□Insert Buffer Bitmap□

page offset 00000002,page type□File Segment inode□

page offset 00000003,page type□B-tree Node□,page level□0001□

page offset 00000004,page type□B-tree Node□,page level□0000□

page offset 00000005,page type□B-tree Node□,page level□0000□

page offset 00000006,page type□B-tree Node□,page level□0000□

page offset 00000007,page type□B-tree Node□,page level□0000□

page offset 00000000,page type□Freshly Allocated Page□

Total number of page:9:

Freshly Allocated Page:1

Insert Buffer Bitmap:1

File Space Header:1

B-tree Node:5
```

```
File Segment inode:1
```

[illegible]page offset[illegible]4[illegible]hexdump[illegible]

```
00010000 b9 aa 8e d0 00 00 00 04 ff ff ff ff ff ff ff ff|................|

00010010 00 00 00 0a ec ea 4e 27 45 bf 00 00 00 00 00 00|......N'E.......|

00010020 00 00 00 00 01 02 00 02 00 ac 80 06 00 00 00 00|................|

00010030 00 a4 00 01 00 03 00 04 00 00 00 00 00 52 d4 8b|.............R..|

00010040 00 00 00 00 00 00 00 00 01 f2 00 00 01 02 00 00|................|

00010050 00 02 02 72 00 00 01 02 00 00 00 02 01 b2 01 00|...r............|

00010060 02 00 41 69 6e 66 69 6d 75 6d 00 05 00 0b 00 00|..Ainfimum......|

00010070 73 75 70 72 65 6d 75 6d 00 00 10 ff f3 7f ff ff|supremum........|

00010080 ff 80 00 00 01 00 00 18 ff f3 7f ff ff fe 80 00|................|

00010090 00 02 00 00 20 ff f3 7f ff ff fd 80 00 00 03 00|..............|

000100a0 00 28 ff f3 7f ff ff fc 80 00 00 04 00 00 00 00|.□.............|

……

00013ff0 00 00 00 00 00 70 00 63 f3 46 77 f2 ec ea 4e 27|.....p.c.Fw...N'|
```

[illegible]4[illegible]c[illegible]4[illegible]5-16[illegible]


辅助索引idx_c
Page Offset:00 04
Key:7f ff ff ff
Pointer:80 00 00 01
Key:7f ff ff fe
Pointer:80 00 00 02
Key:7f ff ff fd
Pointer:80 00 00 03
Key:7f ff ff fc
Pointer:80 00 00 04
聚集索引Page Offset:00 03
Key:80 00 00 01
Pointer:00 05
Key:80 00 00 02
Pointer:00 06
Key:80 00 00 04
Pointer:00 07
Page Offset:00 05
1,repeat（‘a’,7000）,–1
Page Offset:00 06
2,repeat（‘a’,7000）,–2
3,repeat（‘a’,7000）,–3
Page Offset:00 07
4,repeat（‘a’,7000）,–4

[illegible] 5-16 [illegible]

[illegible]5-16[illegible]t[illegible]idx_c[illegible]c[illegible]-1[illegible]7f ff ff ff[illegible]7[illegible]0111[illegible]0[illegible]1[illegible]-1[illegible]

5.4.3　B+□□□□□□

□5.3□□□□B+□□□□□□□□□□□□□□□□□□□□□B+□□□□□□□□□□□□□□□5.3 □□□□□□□□□□□□□B+□□□□□□□□□□□□□

B+□□□

1□2□3□4□5□6□7□8□9

□□□□□□□□□□□□□□□□□□10□□□□□□□□□□□□□□□□□□□□5.3.1□□□□□□□ □□□□□□5□□□□□□□□split record□□□□□□□□□□□□□

P1□1□2□3□4

P2□5□6□7□8□9□10

□□□□□□□□□□□P1□□□□□□□□□□□□□□□□□□□□□□□□□□P2□□□□□□□□□

InnoDB□□□□□Page Header□□□□□□□□□□□□□□□□□□□□

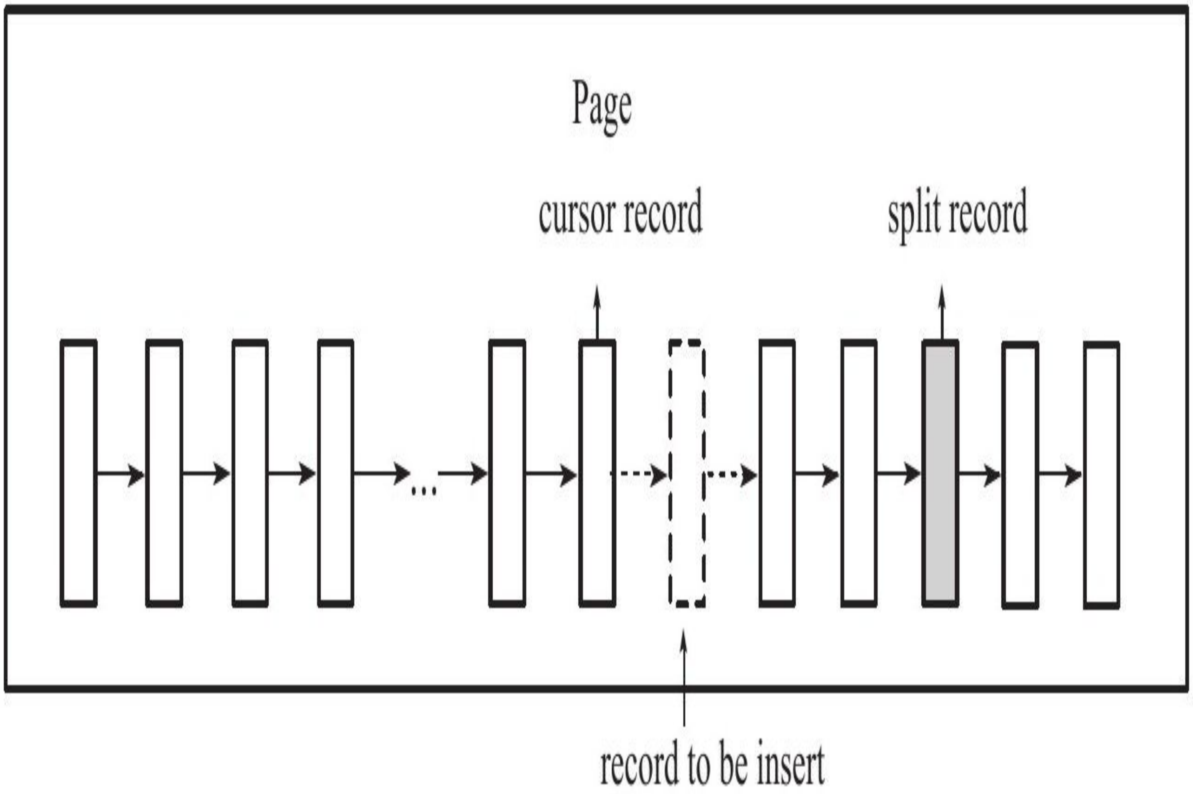
❑PAGE_LAST_INSERT

❑PAGE_DIRECTION

❑PAGE_N_DIRECTION

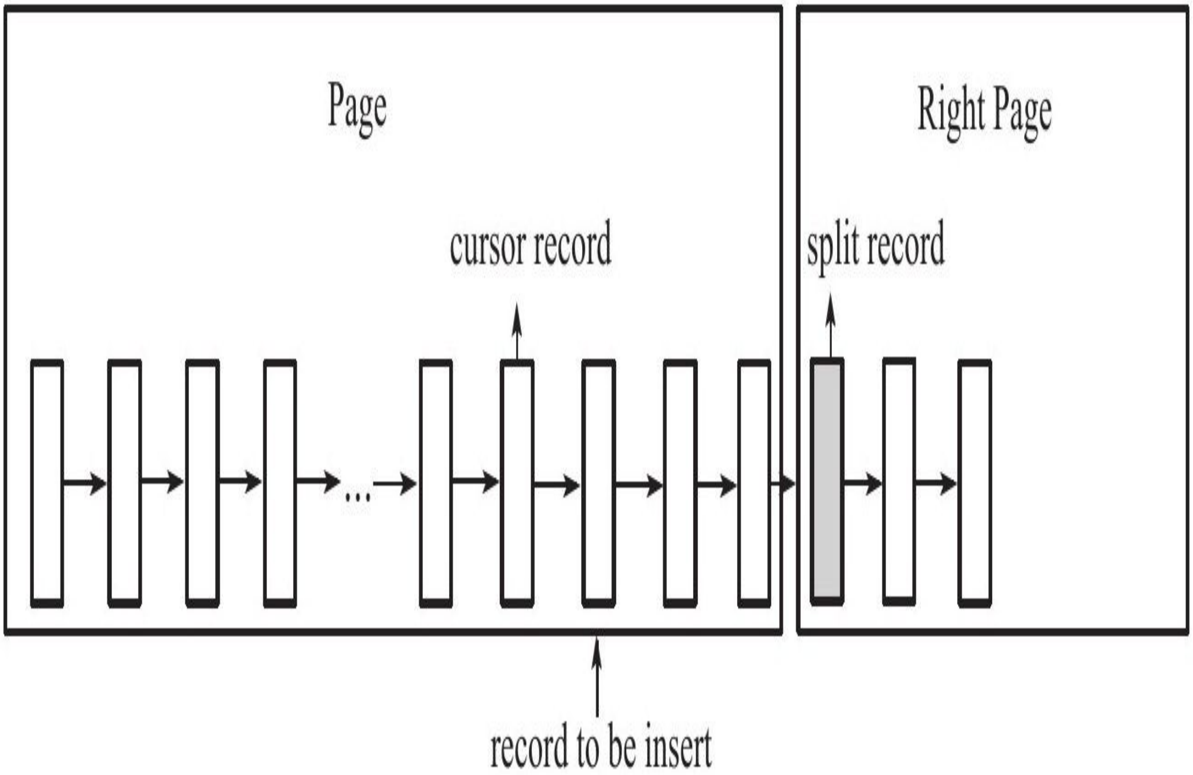
□□□□□□□InnoDB□□5□□□□□□□□□□cursor□□□□□□InnoDB□□3□□

□□□□□□□□□□□□□□□□□□□□□□□□3□□□□□□□□□□□□5-17□□□



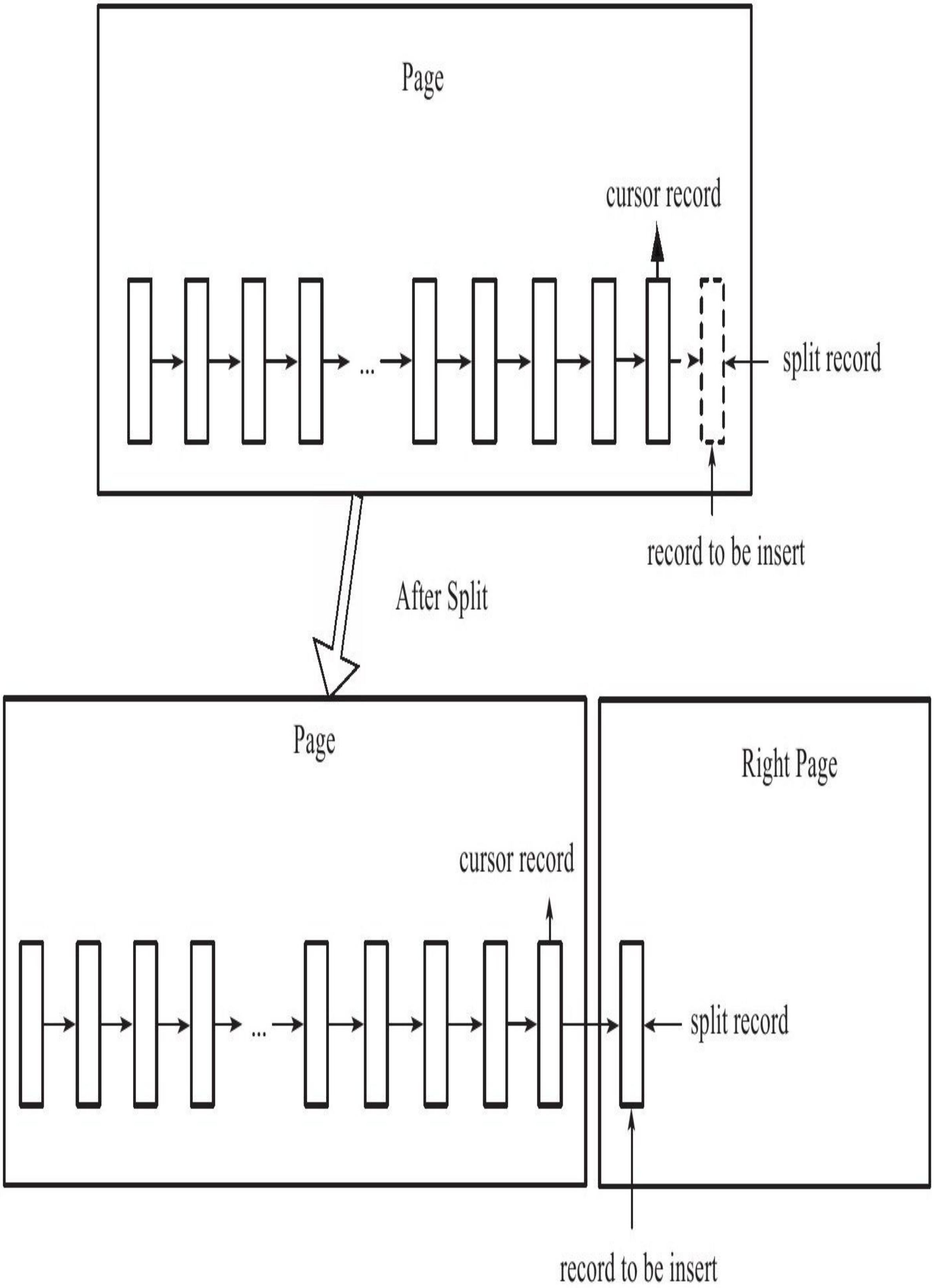
[illegible] 5-17 [illegible]

[illegible]5-17[illegible]3[illegible]split record[illegible]5-18[illegible]


Page
Right Page
cursor record
split record
...
record to be insert

[illegible] 5-18 [illegible]

[illegible]5-19[illegible]


Page
cursor record
split record
record to be insert
After Split
Page
Right Page
cursor record
split record
record to be insert

[illegible] 5-19 [illegible]

5.4.4 B+□□□□□□

1.□□□□

□□□□□□□□□□□□□□□□□□□□ALTER TABLE□□□□□CREATE/DROP INDEX□□□ALTER TABLE□□□□□□□□□

```
ALTER TABLE tbl_name

|ADD{INDEX|KEY}[index_name]

[index_type](index_col_name,...)[index_option]...

ALTER TABLE tbl_name

DROP PRIMARY KEY

|DROP{INDEX|KEY}index_name
```

CREATE/DROP INDEX□□□□□□□□□

```
CREATE[UNIQUE]INDEX index_name

[index_type]

ON tbl_name(index_col_name,...)

DROP INDEX index_name ON tbl_name
```

□□t□□b□varchar□8000□□□□□□□□□□□□100□□□□□□

```
mysql□ALTER TABLE t

-□ADD KEY idx_b(b(100));

Query OK,4 rows affected(0.32 sec)

Records:4 Duplicates:0 Warnings:0
```

[illegible]SHOW INDEX[illegible]t[illegible]a[illegible]c[illegible]idx_a_c[illegible]

```
mysql[illegible]ALTER TABLE t

-[illegible]ADD KEY idx_a_c(a,c);

Query OK,4 rows affected(0.31 sec)

Records:4 Duplicates:0 Warnings:0

mysql[illegible]SHOW INDEX FROM t\G;

***************************1.row***************************

Table:t

Non_unique:0

Key_name:PRIMARY

Seq_in_index:1

Column_name:a

Collation:A

Cardinality:2

Sub_part:NULL

Packed:NULL

Null:

Index_type:BTREE

Comment:

***************************2.row***************************

Table:t

Non_unique:1

Key_name:idx_b

Seq_in_index:1

Column_name:b

Collation:A

Cardinality:2

Sub_part:100
```

```
Packed:NULL
Null:YES
Index_type:BTREE
Comment:
***************************3.row***************************
Table:t
Non_unique:1
Key_name:idx_a_c
Seq_in_index:1
Column_name:a
Collation:A
Cardinality:2
Sub_part:NULL
Packed:NULL
Null:
Index_type:BTREE
Comment:
***************************4.row***************************
Table:t
Non_unique:1
Key_name:idx_a_c
Seq_in_index:2
Column_name:c
Collation:A
Cardinality:2
Sub_part:NULL
Packed:NULL
Null:
Index_type:BTREE
Comment:
```

```
**************************5.row**************************

Table:t

Non_unique:1

Key_name:idx_c

Seq_in_index:1

Column_name:c

Collation:A

Cardinality:2

Sub_part:NULL

Packed:NULL

Null:

Index_type:BTREE

Comment:

5 rows in set[illegible]0.00 sec[illegible]
```

[illegible]SHOW INDEX FROM[illegible]t[illegible]4[illegible]c[illegible]b[illegible]100[illegible]a[illegible]c[illegible]SHOW INDEX[illegible]

❑Table[illegible]

❑Non_unique[illegible]primary key[illegible]0[illegible]

❑Key_name[illegible]DROP INDEX[illegible]

❑Seq_in_index[illegible]idx_a_c[illegible]

❑Column_name[illegible]

❑Collation[illegible]A[illegible]NULL[illegible]B+[illegible]A[illegible]Heap[illegible]Hash[illegible]NULL[illegible]Hash[illegible]Hash[illegible]

❏Cardinality[illegible]Cardinality[illegible]1[illegible]

❏Sub_part[illegible]idx_b[illegible]100[illegible]b[illegible]100[illegible]NULL[illegible]

❏Packed[illegible]NULL[illegible]

❏Null[illegible]NULL[illegible]idx_b[illegible]Yes[illegible]b[illegible]NULL[illegible]

❏Index_type[illegible]InnoDB[illegible]B+[illegible]BTREE[illegible]

❏Comment[illegible]

Cardinality[illegible]Cardinality[illegible]2[illegible]4[illegible]4[illegible]Cardinality[illegible]ANALYZE TABLE[illegible]

```
mysql[illegible]analyze table t\G;

***************************1.row***************************

Table:mytest.t

Op:analyze

Msg_type:status

Msg_text:OK

1 row in set(0.01 sec)

mysql[illegible]show index from t\G;

***************************1.row***************************

Table:t

Non_unique:0

Key_name:PRIMARY

Seq_in_index:1

Column_name:a
```

```
Collation:A

Cardinality:4

Sub_part:NULL

Packed:NULL

Null:

Index_type:BTREE

Comment:

……
```

[illegible]Cardinality[illegible]ANALYZE TABLE[illegible]MySQL[illegible]Cardinality[illegible]

```
mysql[illegible]show index from Profile\G;

***************************1.row***************************

Table:Profile

Non_unique:0

Key_name:UserName

Seq_in_index:1

Column_name:username

Collation:A

Cardinality:NULL

Sub_part:NULL

Packed:NULL

Null:

Index_type:BTREE

Comment:
```

Cardinality[illegible]NULL[illegible]EXPLAIN[illegible]

[illegible]ANALYZE TABLE[illegible]ANALYZE TABLE[illegible]

2.Fast Index Creation

MySQL 5.5[illegible]5.5[illegible]MySQL[illegible]DDL[illegible]MySQL[illegible]

❑[illegible]ALTER TABLE[illegible]

❑[illegible]

❑[illegible]

❑[illegible]

[illegible]Microsoft SQL Server[illegible]Oracle[illegible]DBA[illegible]MySQL[illegible]

InnoDB[illegible]InnoDB 1.0.x[illegible]Fast Index Creation[illegible]——[illegible]FIC[illegible]

[illegible]InnoDB[illegible]S[illegible]InnoDB[illegible]MySQL[illegible]

[illegible]tmpdir[illegible]tmpdir[illegible]

[illegible]FIC[illegible]S[illegible]FIC[illegible]

3.Online Schema Change

Online Schema Change□□□□□□□□□□OSC□□□□□Facebook□□□□□□□□□DDL□□□□□□□□□□□□Facebook□MySQL□□□□□“□□”□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□MySQL□□□□DDL□□□□□□□□

Facebook□□PHP□□□□□OSC□□□□□□□□□InnoDB□□□□□□□□□□OSC□□□Facebook□□□Vamsi Ponnekanti□□□□□□OSC□□□□□□□□□□□□The openarkkit toolkit oak-online-alter-table□□□OSC□□□□□

❑init□□

❑createCopyTable□□□□□□□□□□□□□□□

❑alterCopyTable□□□□□□□□□ALTER TABLE□□□□□□□□□□□□

❑createDeltasTable□□□deltas□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□DML□□□□□□□createDeltasTable□□

❑createTriggers□□□□□□INSERT□UPDATE□DELETE□□□□□□□□□□□□□□□□□□□□deltas□□

❑startSnpshotXact□□□OSC□□□□□□

❑selectTableIntoOutfile□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□chunked□□□□□□□□□□□□□□□□□□□□□□□□□□□□copy□□□□□□□□□□□□□□□□□500 000□

❑dropNCIndexs□□□□□□□□□□□□□□□□□□□□□□□

❑loadCopyTable□□□□□□□□□□□□□□□

❑replayChanges□□OSC□□□□□DML□□□□□□□□□□□□□□□□□□□□□deltas□□□

❑recreateNCIndexes□□□□□□□□□□

❑replayChanges□□□□□DML□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

❑swapTables[illegible]2[illegible]

[illegible]OSC[illegible]OSC[illegible]PHP[illegible]2200[illegible]MySQL InnoDB[illegible]DBA[illegible]InnoDB[illegible]

[illegible]OSC[illegible]PHP[illegible]OSC[illegible]SET sql_bin_log=0[illegible]slave[illegible]

4.Online DDL

[illegible]FIC[illegible]InnoDB[illegible]DML[illegible]OSC[illegible]MySQL 5.6[illegible]Online DDL[illegible]INSERT[illegible]UPDATE[illegible]DELETE[illegible]DML[illegible]MySQL[illegible]

[illegible]DDL[illegible]“[illegible]”[illegible]

❑[illegible]

❑[illegible]

❑[illegible]

❑[illegible]

[illegible]ALTER TABLE[illegible]

```
ALTER  TABLE tbl_name

|ADD{INDEX|KEY}[index_name]

[index_type](index_col_name,...)[index_option]...

ALGORITHM[=]{DEFAULT|INPLACE|COPY}

LOCK[=]{DEFAULT|NONE|SHARED|EXCLUSIVE}
```

ALGORITHM□□□□□□□□□□□□□□COPY□□□□MySQL 5.1□□□□□□□□□□□□□□□□□□□□INPLACE□□□□□□□□□□□□□□□□□□□□DEFAULT□□□□□□old_alter_table□□□□□□INPLACE□□COPY□□□□□□□□□□□□OFF□□□□□INPLACE□□□□□□

```
mysql□SELECT@@version\G;

***************************1.row***************************

@@version:5.6.6-m9

1 row in set(0.00 sec)

mysql□SHOW VARIABLES LIKE'old_alter_table'\G;

***************************1.row***************************Variable_name:old_alter_table

Value:OFF

1 row in set(0.00 sec)
```

LOCK□□□□□□□□□□□□□□□□□□□□□□□□□□□

□1□NONE

□□

□2□SHARE

□□□□□FIC□□□□□□□□□□□□□□□□□□□□□□□□S□□SHARE□□□□□□□□□□□□□

□3□EXCLUSIVE

□EXCLUSIVE□□□□□□□□□□□□□□□□□□□□□□□□□X□□□□□□□□□□□□□□□□□□□□□□□□□COPY□□□□□□□□□□□□□□□□□□COPY□□□□□□□□□□□□

□4□DEFAULT

DEFAULT[illegible]NONE[illegible]SHARE[illegible]EXCLUSIVE[illegible]DEFAULT[illegible]DDL[illegible]

InnoDB[illegible]Online DDL[illegible]INSERT[illegible]UPDATE[illegible]DELETE[illegible]DML[illegible]innodb_online_alter_log_max_size[illegible]128MB[illegible]innodb_online_alter_log_max_size[illegible]

```
Error:1799SQLSTATE:HY000(ER_INNODB_ONLINE_LOG_TOO_BIG)

Message:Creating index'idx_aaa'required more than'innodb_online_alter_log_max_size'bytes of modification log.Please try again.
```

[illegible]innodb_online_alter_log_max_size[illegible]ALTER TABLE[illegible]SHARE[illegible]DML[illegible]

[illegible]Online DDL[illegible]SQL[illegible]

5.5 Cardinality[illegible]

5.5.1 [illegible]Cardinality

[illegible]B+[illegible]B+[illegible]

```
SELECT*FROM student WHERE sex='M'
```

[illegible]'M'[illegible]'F'[illegible]SQL[illegible]50%[illegible]1:1[illegible]B+[illegible]B+[illegible]

[illegible]SHOW INDEX[illegible]Cardinality[illegible]Cardinality[illegible]Cardinality[illegible]Cardinality/n_rows_in_table[illegible]1[illegible]B+[illegible]

```
SELECT*FROM member WHERE usernick='David'
```

[illegible]member[illegible]500[illegible]usernick[illegible]David[illegible]

```
mysql[illegible]EXPLAIN SELECT*FROM member

-[illegible]WHERE usernick='David'\G;

***************************1.row***************************

id:1
```

```
select_type:SIMPLE

table:member

type:const

possible_keys:usernick

key:usernick

key_len:62

ref:const

rows:1

Extra:

1 row in set(0.00 sec)
```

[illegible]usernick[illegible]SQL[illegible]

5.5.2 InnoDB[illegible]Cardinality[illegible]

[illegible]Cardinality[illegible]Cardinality[illegible]Cardinality[illegible]MySQL[illegible]B+[illegible]Cardinality[illegible]

[illegible]Cardinality[illegible]50G[illegible]Cardinality[illegible]Cardinality[illegible]Sample[illegible]

[illegible]InnoDB[illegible]Cardinality[illegible]INSERT[illegible]UPDATE[illegible]INSERT[illegible]UPDATE[illegible]Cardinality[illegible]InnoDB[illegible]Cardinality[illegible]

- [illegible]1/16[illegible]

- stat_modified_counter[illegible]2 000 000 000[illegible]

[illegible]Cardinality[illegible]1/16[illegible]Cardinality[illegible]InnoDB[illegible]stat_modified_counter[illegible]stat_modified_counter[illegible]2 000 000 000[illegible]Cardinality[illegible]

[illegible]InnoDB[illegible]Cardinality[illegible]InnoDB[illegible]8[illegible]Leaf Page[illegible]

- [illegible]B+[illegible]A[illegible]

- [illegible]B+[illegible]8[illegible]P1[illegible]P2[illegible]…[illegible]P8[illegible]

❑□□□□□□□□Cardinality□□□□□Cardinality=□P1+P2+…+P8□*A/8□

□□□□□□□□□□□□□InnoDB□□□□□□Cardinality□□□□□8□□□□□□□□□□□□□□□□□□□□□□□□□□□□Cardinality□□□□□□□□□□□□8□□□□□□□□□□□□□□□□□□□Cardinality□□□□□□□□□Cardinality□□□□□□□□□□

```
SHOW INDEX FROM OrderDetails
```

□□□□SQL□□□□□MySQL□□□□□Cardinality□□□□□□□□□□□□□□□□□5-20□□□

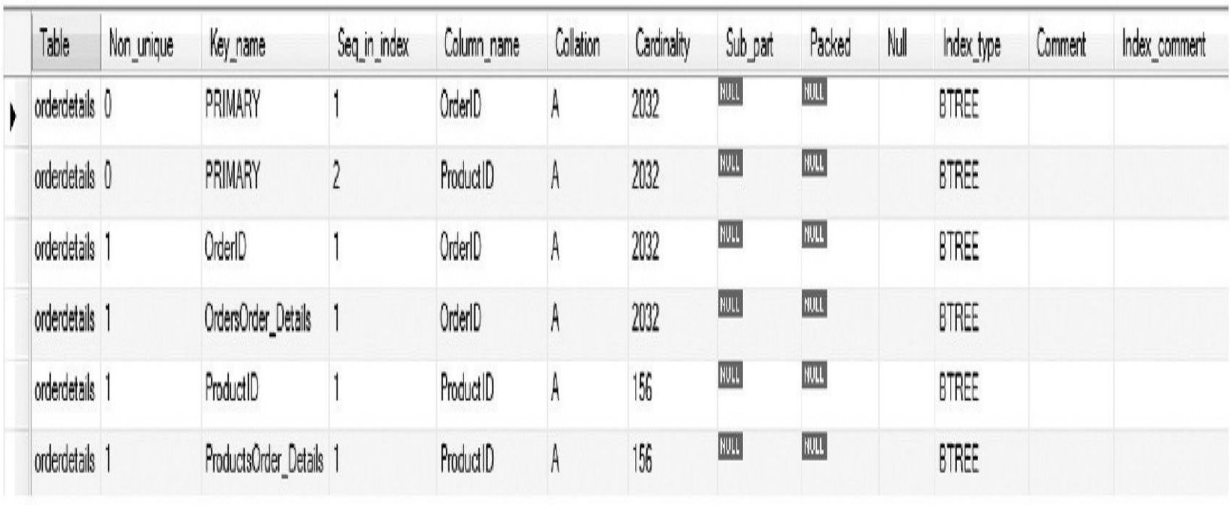


Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment
orderdetails	0	PRIMARY	1	OrderID	A	2032	NULL	NULL		BTREE		
orderdetails	0	PRIMARY	2	ProductID	A	2032	NULL	NULL		BTREE		
orderdetails	1	OrderID	1	OrderID	A	2032	NULL	NULL		BTREE		
orderdetails	1	OrdersOrder_Details	1	OrderID	A	2032	NULL	NULL		BTREE		
orderdetails	1	ProductID	1	ProductID	A	156	NULL	NULL		BTREE		
orderdetails	1	ProductsOrder_Details	1	ProductID	A	156	NULL	NULL		BTREE		

□ 5-20 □□□□□SHOW INDEX FROM OrderDetails□□□

□□□□□□□□□□□□□□INSERT□UPDATE□□□□□□□□OrderDetails□□□□□□□□□□□□□□SHOW INDEX FROM□□□□Cardinality□□□□□□□□□□□5-21□□□

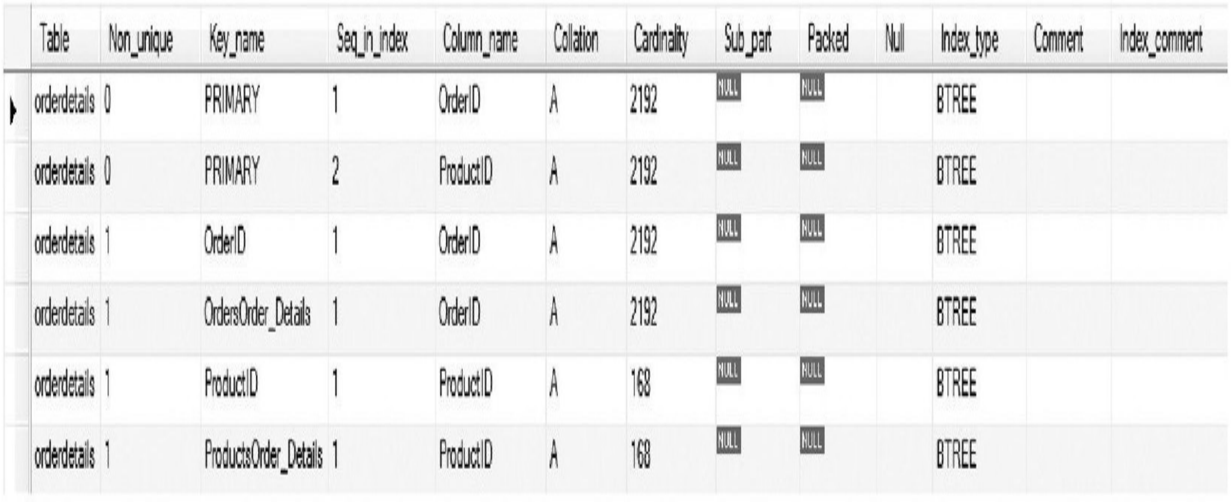


Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment
orderdetails	0	PRIMARY	1	OrderID	A	2192	NULL	NULL		BTREE		
orderdetails	0	PRIMARY	2	ProductID	A	2192	NULL	NULL		BTREE		
orderdetails	1	OrderID	1	OrderID	A	2192	NULL	NULL		BTREE		
orderdetails	1	OrdersOrder_Details	1	OrderID	A	2192	NULL	NULL		BTREE		
orderdetails	1	ProductID	1	ProductID	A	168	NULL	NULL		BTREE		
orderdetails	1	ProductsOrder_Details	1	ProductID	A	168	NULL	NULL		BTREE		

□ 5-21 □□□□□SHOW INDEX FROM OrderDetails□□□

□□□□□□□□□□SHOW INDEX FROM□□□□□OrderDetails□□□□Cardinality□□□□□□□□□□□OrderDetails□□□□□□□□□□□□□□□□□□Cardinality□□□□□8□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Cardinality□□□□□□□□□□□□□□□InnoDB□□□□□Bug□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□Cardinality□□□□□□□□□□□□□□□□□□□□□□□□□□□□8□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Cardinality□□□□□□

□InnoDB 1.2□□□□□□□□□□□innodb_stats_sample_pages□□□□□□Cardinality□□□□□□□□□□□□□□8□□□□□□innodb_stats_method□□□□□□□□□□□□□□NULL□□□□□□□□□□□nulls_equal□□□□NULL□□□□□□□□□□□□□□□□□nulls_unequal□nulls_ignored□□□□□□NULL□□□□□□□□□□□□□NULL□□□□□□□□□□□□□□NULL□NULL□1□2□2□3□3□3□□□□innodb_stats_method□□□□□□□□□□Cardinality□4□□□□innodb_stats_method□nulls_unequal□□□□□Caridinality□5□□□□innodb_stats_method□nulls_ignored□□Cardinality□3□

□□□SQL□□ANALYZE TABLE□SHOW TABLE STATUS□SHOW INDEX□□□□INFORMATION_SCHEMA□□□□□TABLES□STATISTICS□□□□InnoDB□□□□□□□□□□□□Cardinality□□Cardinality□□

InnoDB1.2□□□□□□□□□□□Cardinality□□□□□□□□□□□□□5-3□□□

表 5-3　InnoDB 1.2 新增参数

参数	说明
innodb_stats_persistent	是否将命令 ANALYZE TABLE 计算得到的 Cardinality 值存放到磁盘上。若是，则这样做的好处是可以减少重新计算每个索引的 Cardinality 值，例如当 MySQL 数据库重启时。此外，用户也可以通过命令 CREATE TABLE 和 ALTER TABLE 的选项 STATS_PERSISTENT 来对每张表进行控制。 默认值：OFF
innodb_stats_on_metadata	当通过命令 SHOW TABLE STATUS、SHOW INDEX 及访问 INFORMATION_SCHEMA 架构下的表 TABLES 和 STATISTICS 时，是否需要重新计算索引的 Cardinality 值。 默认值：OFF
innodb_stats_persistent_sample_pages	若参数 innodb_stats_persistent 设置为 ON，该参数表示 ANALYZE TABLE 更新 Cardinality 值时每次采样页的数量。 默认值：20
innodb_stats_transient_sample_pages	该参数用来取代之前版本的参数 innodb_stats_sample_pages，表示每次采样页的数量。 默认值为：8

5.6 B+[illegible]

5.6.1 [illegible]B+[illegible]

[illegible]B+[illegible]B+[illegible]DBA[illegible]Think Different[illegible]

[illegible]1[illegible]OLTP[illegible]OLAP[illegible]OLTP[illegible]10[illegible]1[illegible]OLTP[illegible]B+[illegible]B+[illegible]

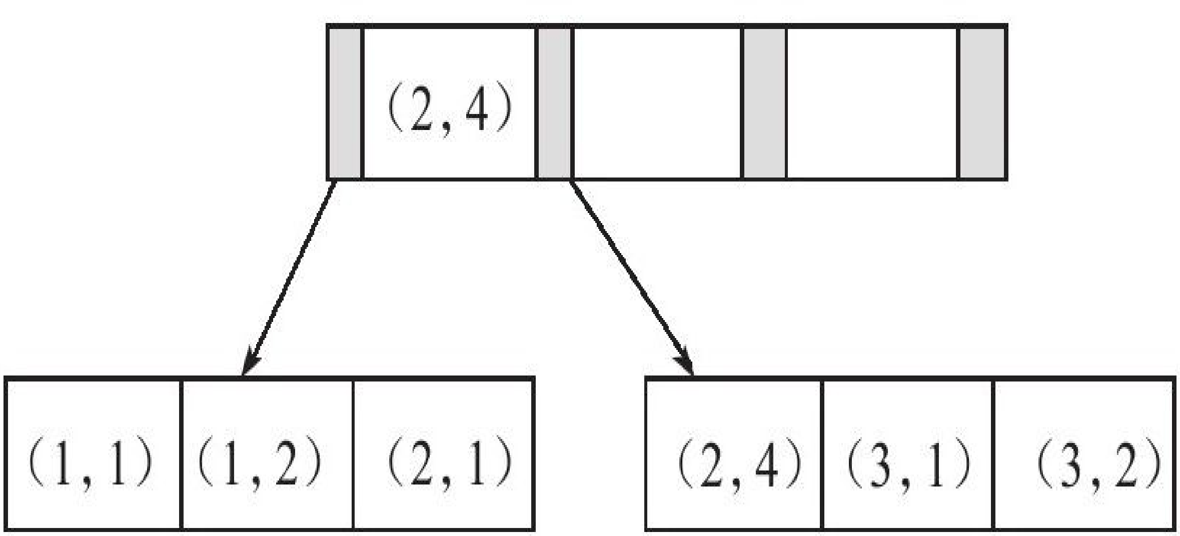
[illegible]OLAP[illegible]OLAP[illegible]OLAP[illegible]OLAP[illegible]OLAP[illegible]Hash Join[illegible]DBA[illegible]OLAP[illegible]

5.6.2 □□□□

□□□t□□□□□□idx_a_b□□□□□□□□□□□□□a□b□□

```
CREATE TABLE t(

a INT,

b INT,

PRIMARY KEY(a),

KEY idx_a_b(a,b)

)ENGINE=INNODB
```

□□□B+□□□□□□□□□□□□□□□□□□1□□□□□□□2□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□a□b□□□□5-22□□□



□ 5-22 □□□□□B+□

[illegible]5-22[illegible]B+[illegible]B+[illegible]1[illegible]1[illegible]1[illegible]2[illegible]2[illegible]1[illegible]2[illegible]4[illegible]3[illegible]1[illegible]3[illegible]2[illegible]a[illegible]b[illegible]

[illegible]SELECT*FROM TABLE WHERE a=xxx and b=xxx[illegible]a[illegible]b[illegible]a[illegible]SELECT*FROM TABLE WHERE a=xxx[illegible]a[illegible]b[illegible]b[illegible]SELECT*FROM TABLE WHERE b=xxx[illegible]B+[illegible]b[illegible]1[illegible]2[illegible]1[illegible]4[illegible]1[illegible]2[illegible]b[illegible]a[illegible]b[illegible]

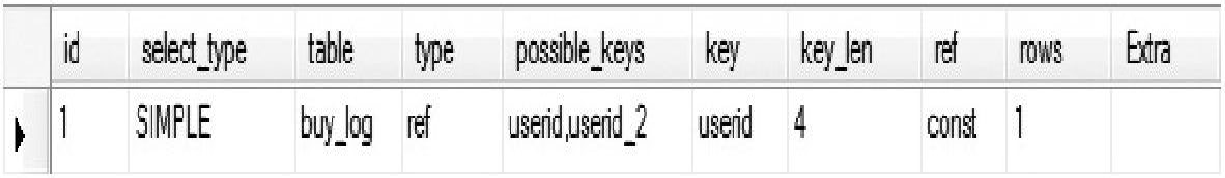
[illegible]buy_log[illegible]

```
CREATE TABLE buy_log(

userid INT UNSIGNED NOT NULL,

buy_date DATE

)ENGINE=InnoDB;

INSERT INTO buy_log VALUES(1,'2009-01-01');

INSERT INTO buy_log VALUES(2,'2009-01-01');

INSERT INTO buy_log VALUES(3,'2009-01-01');

INSERT INTO buy_log VALUES(1,'2009-02-01');

INSERT INTO buy_log VALUES(3,'2009-02-01');

INSERT INTO buy_log VALUES(1,'2009-03-01');

INSERT INTO buy_log VALUES(1,'2009-04-01');

ALTER TABLE buy_log ADD KEY(userid);

ALTER TABLE buy_log ADD KEY(userid,buy_date);
```

[illegible]userid[illegible]userid[illegible]

```
SELECT*FROM buy_log WHERE userid=2;
```

□□□□□□□□□□5-23□□□



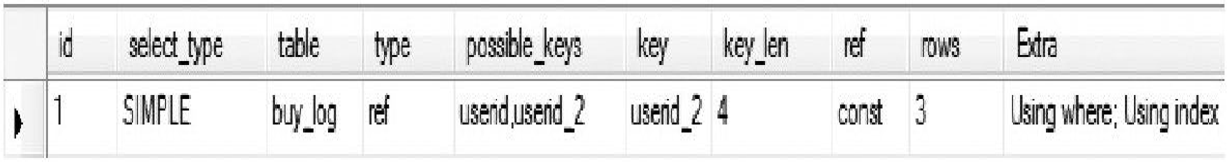
id	select_type	table	type	possible_keys	key	key_len	ref	rows	Extra
1	SIMPLE	buy_log	ref	userid,userid_2	userid	4	const	1	

□   5-23   □□□□□□userid□□□□□

□□5-23□□□□□□possible_keys□□□□□□□□□□□□□□□□□□□userid□□□□userid□buy_date□□□□□□□□□□□□□□□□□□□□userid□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□userid□1□□□3□□□□□□□□SQL□□□□□□□□□□□5-24□□□

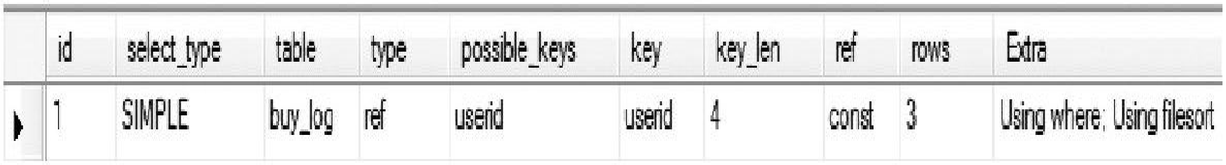
```
SELECT*FROM buy_log

WHERE userid=1 ORDER BY buy_date DESC LIMIT 3
```



id	select_type	table	type	possible_keys	key	key_len	ref	rows	Extra
1	SIMPLE	buy_log	ref	userid,userid_2	userid_2	4	const	3	Using where; Using index

□   5-24   SQL□□□□□□□

□□□□□□□□□SQL□□□□□□□userid□□□□□□□□□userid□buy_date□□□□□□□□□□□□□□□userid□buy_date□□□□□□userid_2□□□□□□□□□□□buy_date□□□□□□□□□□□□□□□□□□□□□□□buy_date□□□□□□□□□□□□□□□□userid□□□□□□□□□□5-25□□□



id	select_type	table	type	possible_keys	key	key_len	ref	rows	Extra
1	SIMPLE	buy_log	ref	userid	userid	4	const	3	Using where; Using filesort

□   5-25   □□□□userid□□□□□□□

□Extra□□□□□□□Using filesort□□□□□□□□□□□□□□□□□□□□□□□□□□□□□buy_date□□□□□□□userid□□buy_date□□□□□□

□□□□□□□□□□□□□□□□a□b□□□□□□□a□b□□□□□□□□□□□□□□□□□□□□□□□□□□

```
SELECT...FROM TABLE WHERE a=xxx ORDER BY b
```

□□□□□□□□□a□b□c□□□□□□□□□□□□□□□□□□□□□□□□□

```
SELECT...FROM TABLE WHERE a=xxx ORDER BY b

SELECT...FROM TABLE WHERE a=xxx AND b=xxx ORDER BY c
```

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□filesort□□□□□□□□□□a□c□□□□□□

```
SELECT...FROM TABLE WHERE a=xxx ORDER BY c
```

5.6.3 □□□□

InnoDB□□□□□□□□□□□covering index□□□IO□□□

□□ □□□□□□□□□□InnoDB Plugin□□□□□□□□□□□□□InnoDB□□□□1.0□□□□MySQL□□□□□□5.0□□□□□InnoDB□□□□□□□□□□□□□□

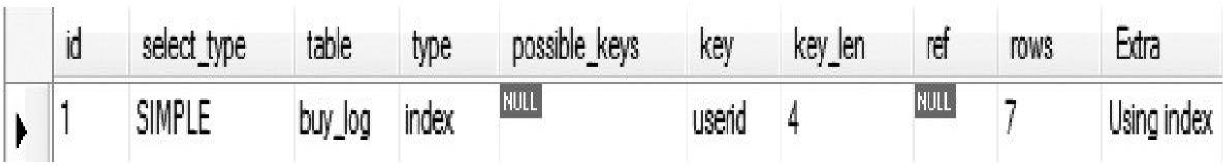
□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□primary key1□primary key2□…□key1□key2□…□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
SELECT key2 FROM table WHERE key1=xxx□

SELECT primary key2,key2 FROM table WHERE key1=xxx□

SELECT primary key1,key2 FROM table WHERE key1=xxx□

SELECT primary key1,primary key2□key2 FROM table WHERE key1=xxx□
```

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□buy_log□□□□□□□□□□

```
SELECT COUNT(*)FROM buy_log;
```

InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□buy_log□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□IO□□□□□□□□□□□□□5-26□□□



	id	select_type	table	type	possible_keys	key	key_len	ref	rows	Extra
▶	1	SIMPLE	buy_log	index	NULL	userid	4	NULL	7	Using index

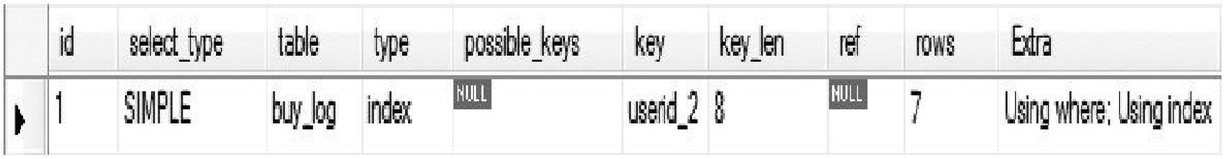
□ 5-26 COUNT□*□□□□□□□□

□□□5-26□□□□□possible_keys□□NULL□□□□□□□□□□□□□□□userid□□□□□Extra□□Using index□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□a□b□□□□□□□□□□□□□□□□b□□

```
SELECT COUNT(*)FROM buy_log

WHERE buy_date□='2011-01-01'AND buy_date□'2011-02-01'
```

□buy_log□□userid□buy_date□□□□□□□□□□□□□b□□□□□□□□□□□□□□□□□□□□□□□□□□□□SQL□□□5-27□□□



id	select_type	table	type	possible_keys	key	key_len	ref	rows	Extra
1	SIMPLE	buy_log	index	NULL	userid_2	8	NULL	7	Using where; Using index

□ 5-27 □□□□□□□□□□□□

□□5-27□□□□□□possible_keys□□□NULL□□□□key□userid_2□□□□□userid□buy_date□□□□□□□□□Extra□□□□□□Using index□□□□□□□□□□□

5.6.4 □□□□□□□□□□□□□

□□□□□□□□□□EXPLAIN□□□□SQL□□JOIN□□□□□□□□□□□□

```
SELECT*FROM orderdetails

WHERE orderid□10000 and orderid□102000;
```

□□□□SQL□□□□□□□□□10000□□□□□□□□□□SHOW INDEX FROM orderdetails□□□□□□□□□□5-28□□□

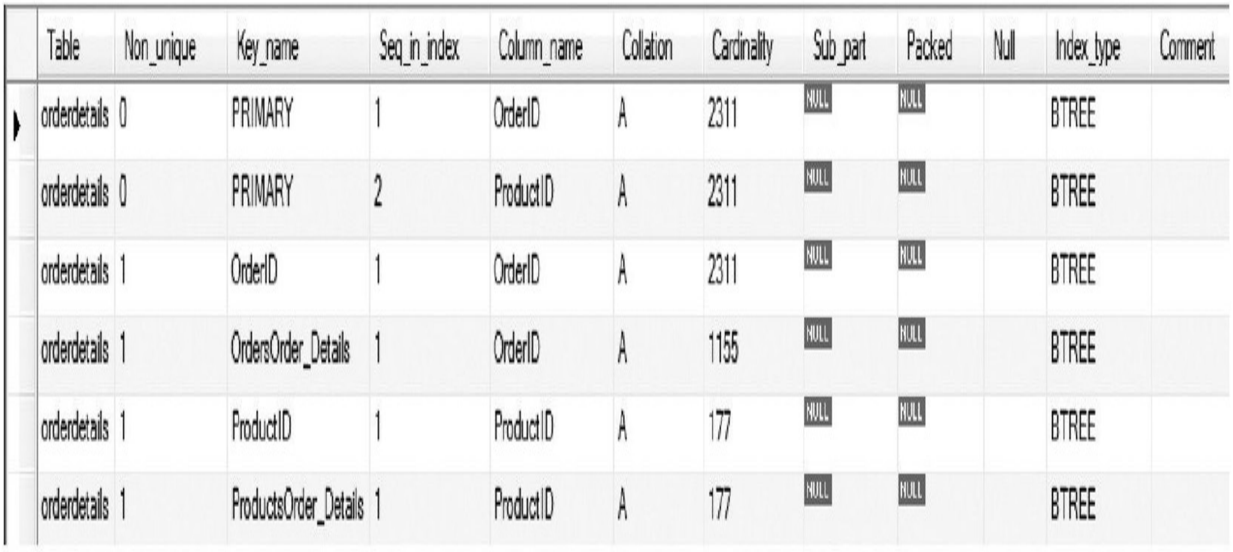
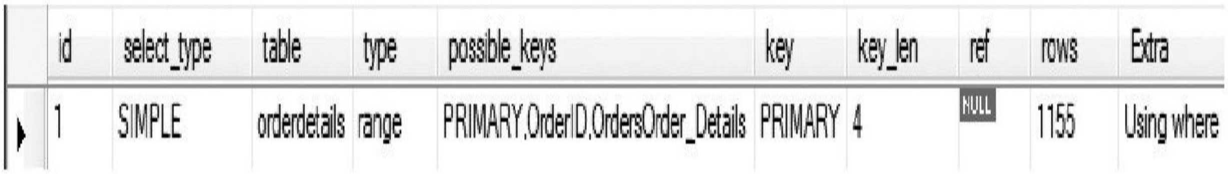


Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment
orderdetails	0	PRIMARY	1	OrderID	A	2311	NULL	NULL		BTREE	
orderdetails	0	PRIMARY	2	ProductID	A	2311	NULL	NULL		BTREE	
orderdetails	1	OrderID	1	OrderID	A	2311	NULL	NULL		BTREE	
orderdetails	1	OrdersOrder_Details	1	OrderID	A	1155	NULL	NULL		BTREE	
orderdetails	1	ProductID	1	ProductID	A	177	NULL	NULL		BTREE	
orderdetails	1	ProductsOrder_Details	1	ProductID	A	177	NULL	NULL		BTREE	

□ 5-28 □orderdetails□□□□□

□□□□□orderdetails□□OrderID□ProductID□□□□□□□□□□□□□□OrderID□□□□□□□□□□SQL□□□□□□□□□OrderID□□□□□□□□□□□□□□□□EXPLAIN□□□□□□□□□□□□□□□□OrderID□□□□□□□□□□□□5-29□□□



id	select_type	table	type	possible_keys	key	key_len	ref	rows	Extra
1	SIMPLE	orderdetails	range	PRIMARY,OrderID,OrdersOrder_Details	PRIMARY	4	NULL	1155	Using where

□ 5-29 □□□□□□□SQL□□□□

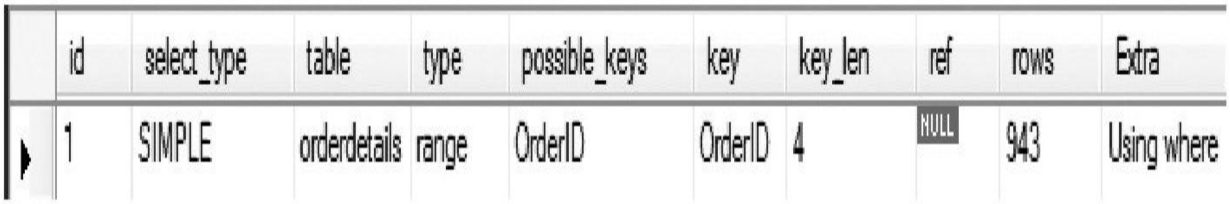
□possible_keys□□□□□□□□□□□□PRIMARY□OrderID□OrdersOrder_Details□□□□□□□□□□□□□□□□□□□□□□□PRIMARY□□□□□□□□□□□□table scan□□□□OrderID□□□□□□□index scan□□

□□□□□□□□□□□□□□□□□□□□□□□□□□OrderID□□□□□□□□□□□□□□□□□□□□OrderID□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□OrderID□□□20%□□□

□□FORCE INDEX□□□□□□□□□□□□

```
SELECT*FROM orderdetails FORCE INDEX(OrderID)

WHERE orderid□10000 and orderid□102000;
```

□□□□□□□□□5-30□□□



id	select_type	table	type	possible_keys	key	key_len	ref	rows	Extra
1	SIMPLE	orderdetails	range	OrderID	OrderID	4	NULL	943	Using where

□ 5-30 □□□□□□□□

5.6.5 □□□□

MySQL□□□□□□□□□□INDEX HINT□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□INDEX HINT□

❑MySQL□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□□□ MySQL□□□ DBA□□□□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□

❑□SQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□Range□□□□□□□□□□□□□□□□DBA□□□□□□□□□□□□□□□□□Index Hint□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□MySQL□□□□Index Hint□□□□□□

```
tbl_name[[AS]alias][index_hint_list]

index_hint_list:

index_hint[,index_hint]...

index_hint:

USE{INDEX|KEY}

[{FOR{JOIN|ORDER BY|GROUP BY}]([index_list])

|IGNORE{INDEX|KEY}

[{FOR{JOIN|ORDER BY|GROUP BY}](index_list)

|FORCE{INDEX|KEY}

[{FOR{JOIN|ORDER BY|GROUP BY}](index_list)

index_list:

index_name[,index_name]...
```

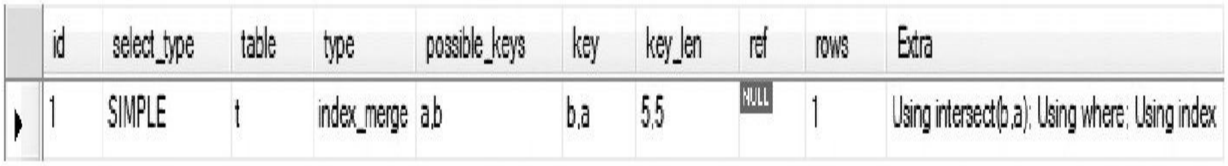
□□□□□□□□□□□□□□□□□□□□□□t□□□□□□□□□

```
CREATE TABLE t(

a INT,
b INT,

KEY(a),

KEY(b)

)ENGINE=INNODB;

INSERT INTO t SELECT 1,1;

INSERT INTO t SELECT 1,2;

INSERT INTO t SELECT 2,3;

INSERT INTO t SELECT 2,4;

INSERT INTO t SELECT 1,2;
```

□□□□□□□SQL□□□

```
SELECT*FROM t WHERE a=1 AND b=2;
```

□□EXPLAIN□□□□□□5-31□□□□□□□□



id	select_type	table	type	possible_keys	key	key_len	ref	rows	Extra
1	SIMPLE	t	index_merge	a,b	b,a	5,5	NULL	1	Using intersect(b,a); Using where; Using index

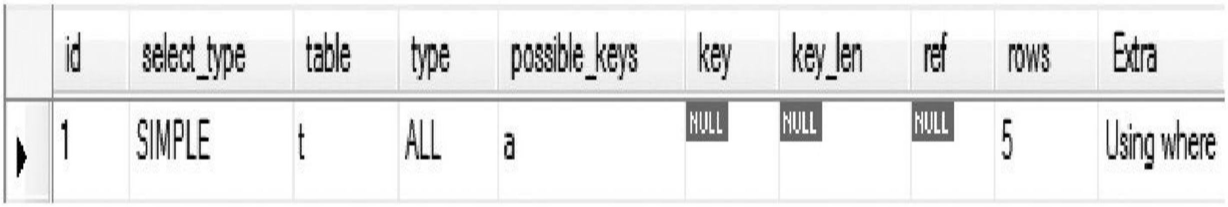
□ 5-31 SQL□□□□□□□

□5-31□□□possible_keys□□□□□SQL□□□□□□□□□a□b□□□□□□□□□□□key□□□□□□a□b□□□□MySQL□□□□□a□b□□□□□□□□□□□□□□Extra□□□Using intersect□b□a□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□USE INDEX□□□□□□□□a□□□□□□□

```
SELECT*FROM t USE INDEX(a)WHERE a=1 AND b=2;
```

□□□□□□□□□5-32□□□



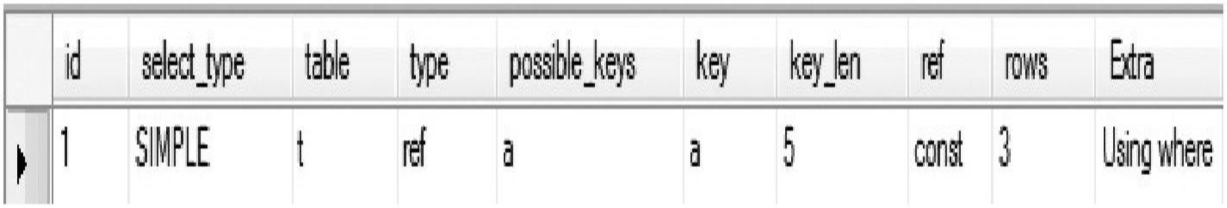
id	select_type	table	type	possible_keys	key	key_len	ref	rows	Extra
1	SIMPLE	t	ALL	a	NULL	NULL	NULL	5	Using where

□ 5-32 □□USE INDEX□□□□□□

□□□□□□□□□□□□□a□□□□□□□□□□□□□□□□□□□□□□□□□□USE INDEX□□□FORCE INDEX□□□□□□□□

```
SELECT*FROM t FORCE INDEX(a)WHERE a=1 AND b=2;
```

□□□□□□□□□□5-33□□□



id	select_type	table	type	possible_keys	key	key_len	ref	rows	Extra
1	SIMPLE	t	ref	a	a	5	const	3	Using where

□ 5-33 □□FORCE INDEX□□□□□□

□□□FORCE INDEX□□□□USE INDEX□

5.6.6 Multi-Range Read[illegible]

MySQL5.6[illegible]Multi-Range Read[illegible]MRR[illegible]Multi-Range Read[illegible]IO-bound[illegible]SQL[illegible]Multi-Range Read[illegible]range[illegible]ref[illegible]eq_ref[illegible]

MRR[illegible]

❑MRR[illegible]

❑[illegible]

❑[illegible]

[illegible]InnoDB[illegible]MyISAM[illegible]JOIN[illegible]MRR[illegible]

❑[illegible]

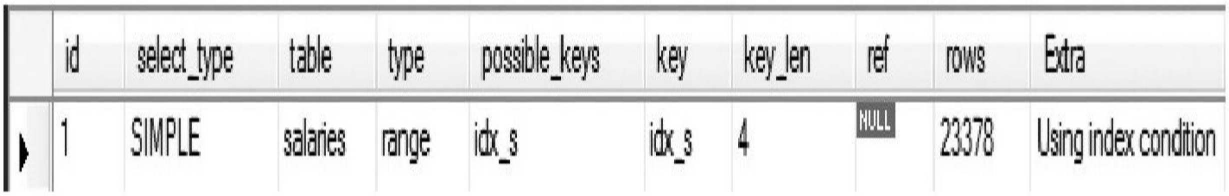
❑[illegible]RowID[illegible]

❑[illegible]RowID[illegible]

[illegible]InnoDB[illegible]MyISAM[illegible]SQL[illegible]

```
SELECT*FROM salaries WHERE salary[illegible]10000 AND salary[illegible]40000;
```

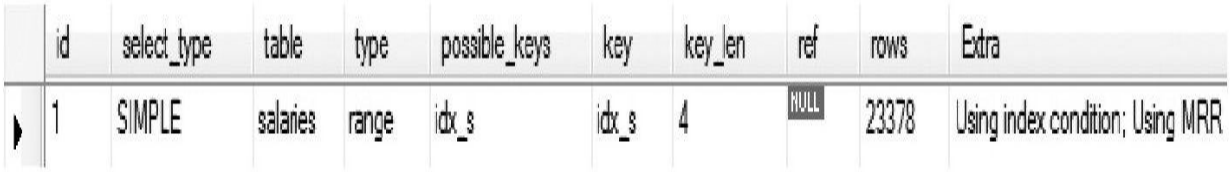
salary[illegible]idx_s[illegible]Multi-Range Read[illegible]5-34[illegible]



id	select_type	table	type	possible_keys	key	key_len	ref	rows	Extra
1	SIMPLE	salaries	range	idx_s	idx_s	4	NULL	23378	Using index condition

图 5-34 □□□Multi-Range Read□□□□□

□□□Mulit-Range Read□□□□□□□□□Extra□□Using index condition □□□□□□Using MRR□□□□□5-35□□□

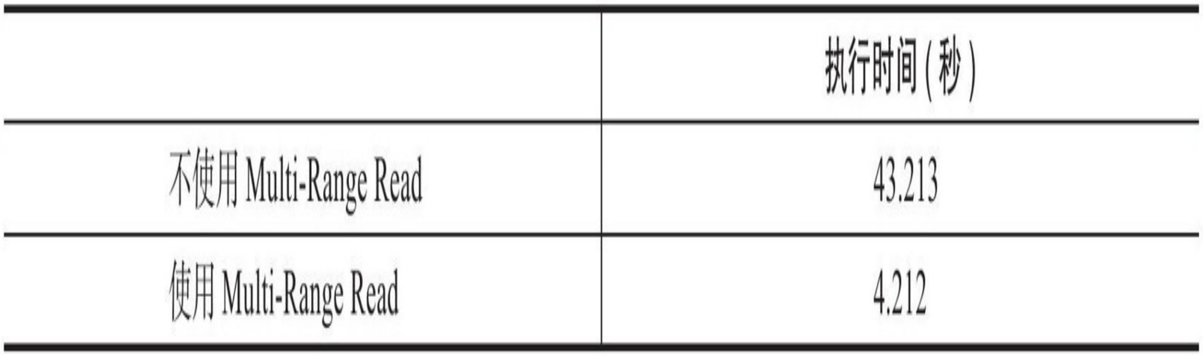


id	select_type	table	type	possible_keys	key	key_len	ref	rows	Extra
1	SIMPLE	salaries	range	idx_s	idx_s	4	NULL	23378	Using index condition; Using MRR

□ 5-35 □□Multi-Range Read□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□5-4□□□

表 5-4 是否启用 Multi-Range Read 的执行时间对比



	执行时间(秒)
不使用 Multi-Range Read	43.213
使用 Multi-Range Read	4.212

□□□□□□□□□□□□□□□□□□□□□□□10□□□□□□Multi-Range Read□□□□□□□□□□□□□□□□□□□□

□□ □□□□□□□MySQL□□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□Multi-Range Read□□

```
SELECT*FROM t

WHERE key_part1□=1000 AND key_part1□2000

AND key_part2=10000;
```

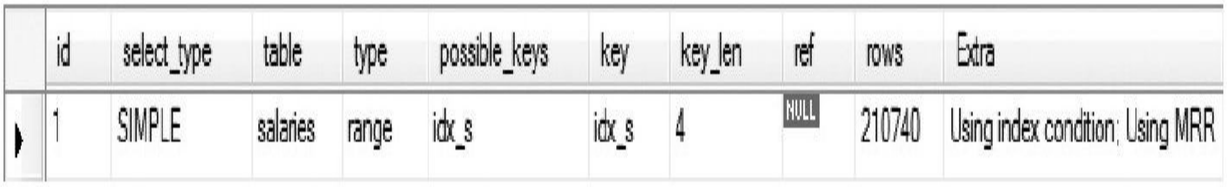
□t□□key_part1□key_part2□□□□□□□□□□□□□key_part1□key_part2□□□□□□□□□□□□□Multi-Read Range□□□□□□□□Range□SQL□□□□□□key_part1□□1000□□□2000□□□□□□□□□key_part2□□□1000□□□□□□□□□□□key_part2□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□key_part2□□□1000□□□□Mulit-Range Read□□□□□□□□□□□□□

□□□□□Multi-Range Read□□□1000□1000□□□1001□1000□□□1002□1000□□…□□1999□1000□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□

```
SELECT*FROM salaries

WHERE(from_date between'1986-01-01'AND'1995-01-01')

AND(salary between 38000 and 40000);
```

□□□□Multi-Range Read□□□□□□□□□□5-36□□□



id	select_type	table	type	possible_keys	key	key_len	ref	rows	Extra
1	SIMPLE	salaries	range	idx_s	idx_s	4	NULL	210740	Using index condition; Using MRR

□ 5-36 □□Multi-Range Read□□□□□

□salaries□□□□salary□□□idx_s□□□□□□SQL□□□□□□□□□Multi-Range Read□□□□□□□□□□□□□□□□□□□□Extra□□□□□Using MRR□□□

□□□□Multi-Range Read□□□□□□□□optimizer_switch□□□flag□□□□□□mrr□on□□□□□□Multi-Range Read□□□mrr_cost_based□□□□□□□□cost based□□□□□□□□□□mrr□□□mrr□□on□mrr_cost_based□□off□□□□□□Multi-Range Read□□□□□□□□□□□□□Multi-Range Read□□□□□□□□□□□

```
mysql□SET@@optimizer_switch='mrr=on,mrr_cost_based=off';

Query OK,0 rows affected(0.00 sec)
```

□□read_rnd_buffer_size□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□RowID□□□□□□□□RowID□□□□□□□□□□□256K□

```
mysql□SELECT@@read_rnd_buffer_size\G;

***************************1.row***************************

@@read_rnd_buffer_size:262144

1 row in set(0.00 sec)
```

5.6.7 Index Condition Pushdown□ICP□□□

□Multi-Range Read□□□Index Condition Pushdown□□□MySQL 5.6□□□□□□□□□□□□□□□□□□□□□□□□□MySQL□□□□□□□□Index Condition Pushdown□□□□□□□□□□□□□□□□□□□□□□□□□□□□WHERE□□□□□□□□□□□Index Condition Pushdown□□MySQL□□□□□□□□□□□□□□□□□□□□□□WHERE□□□□□□□□□□WHERE□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□□fetch□□□□□□□□□□□□□□□□

Index Condition Pushdown□□□□range□ref□eq_ref□ref_or_null□□□□□□□□□□MyISAM□InnoDB□□□□□□□□□□□Index Condition Pushdown□□□□□□□□□□□□Extra□□Using index condition□□□

□□ NDB Cluster□□□□□□Engine Condition Pushdown□□□□□□□□□“Index”□Condition Pushdown□□□□□□□□□□Condition Pushdown□□□□□□□□□□□□□□□□□□□□□□MySQL 5.1□□□NDB Cluster□□□□□□□□□Engine Condition Pushdown□□□

□□□□□□□□□□(zip_code□last_name□firset_name)□□□□□□□□□□

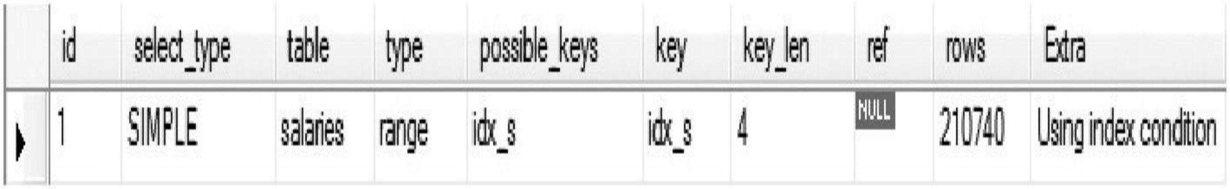
```
SELECT*FROM people

WHERE zipcode='95054'

AND lastname LIKE'%etrunia%'

AND address LIKE'%Main Street%';
```

□□□□□□□MySQL□□□□□□□□□□□□zipcode□□95 054□□□□□□□□□WHERE□□□lastname LIKE'%etrunia%'AND address LIKE'%Main Street%'□□□□□□□□□□□Index Condition Pushdown□□□□□□□□□□□□□□□□□□zipcode□□95 054□□□□□□□□□WHERE□□□□□□□□

□□□Index Condition Pushdown□□□□□□□□□□□□□□□WHERE□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□WHERE□□□□□□□□□□□□□□□□□□□□□□□□□□SQL□□□

```
SELECT*FROM salaries

WHERE(from_date between'1986-01-01'AND'1995-01-01')

AND(salary between 38000 and 40000);
```

□□□□Multi-Range Read□□□□□□□□□□□5-37□□□



id	select_type	table	type	possible_keys	key	key_len	ref	rows	Extra
1	SIMPLE	salaries	range	idx_s	idx_s	4	NULL	210740	Using index condition

□ 5-37 □□□Multi-Range Read□□□□□□□

□□□□□Extra□Using index condition□□□□□□□□□□□□idx_s□□□□□ Index Condition Pushdown□□□□□□□□□□□□□(emp_no□ from_date)□□□□□□□□idx_s□□□□□□from_date□□□□□□□□□□□□□□

□5-5□□□□MySQL 5.5□MySQL 5.6□□□SQL□□□□□□□□□□□□□□□□ MRR□□□□□□□

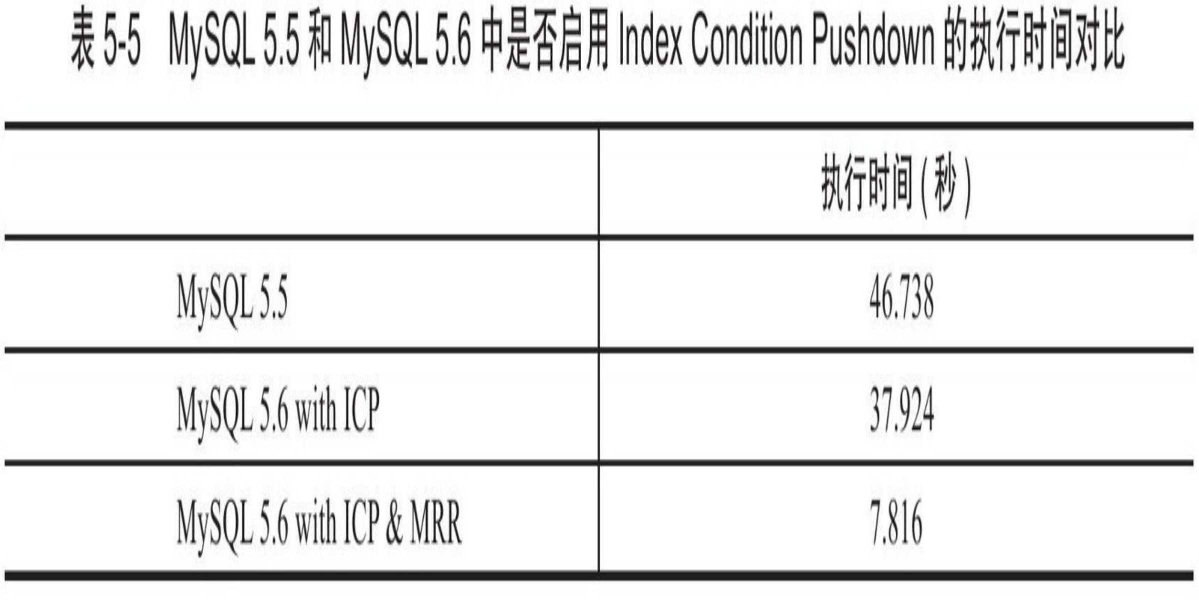


表 5-5 MySQL 5.5 和 MySQL 5.6 中是否启用 Index Condition Pushdown 的执行时间对比

	执行时间（秒）
MySQL 5.5	46.738
MySQL 5.6 with ICP	37.924
MySQL 5.6 with ICP & MRR	7.816

[illegible]Index Condition Pushdown[illegible]MySQL 5.5[illegible]23%[illegible]Mulit-Range Read[illegible]400%[illegible]

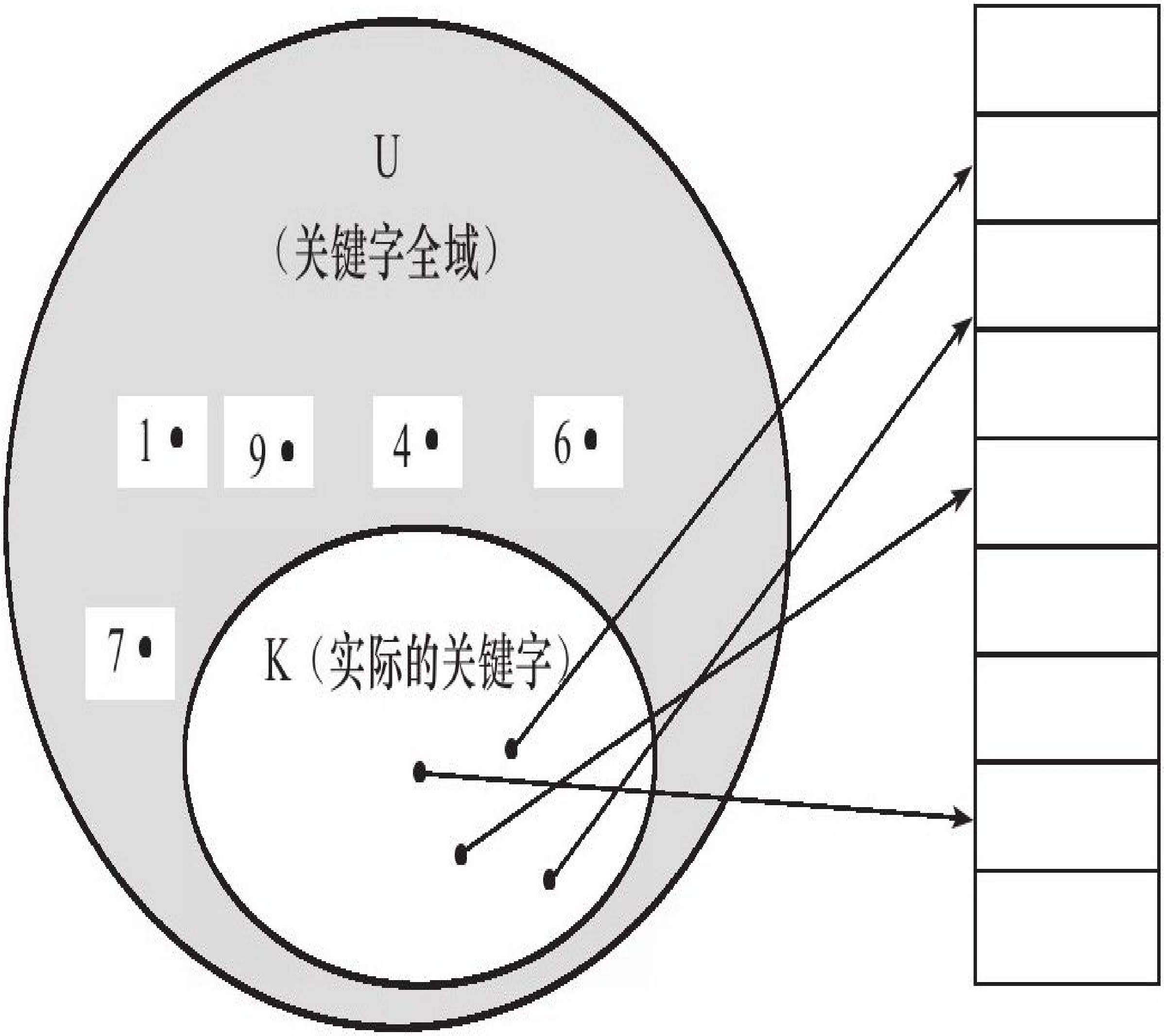
5.7 [illegible]

[illegible]O[illegible]1[illegible]128GB[illegible]O[illegible]1[illegible]

5.7.1 [illegible]

[illegible]Hash Table[illegible]U[illegible]U={0[illegible]1[illegible]…[illegible]m-1} [1] [illegible]

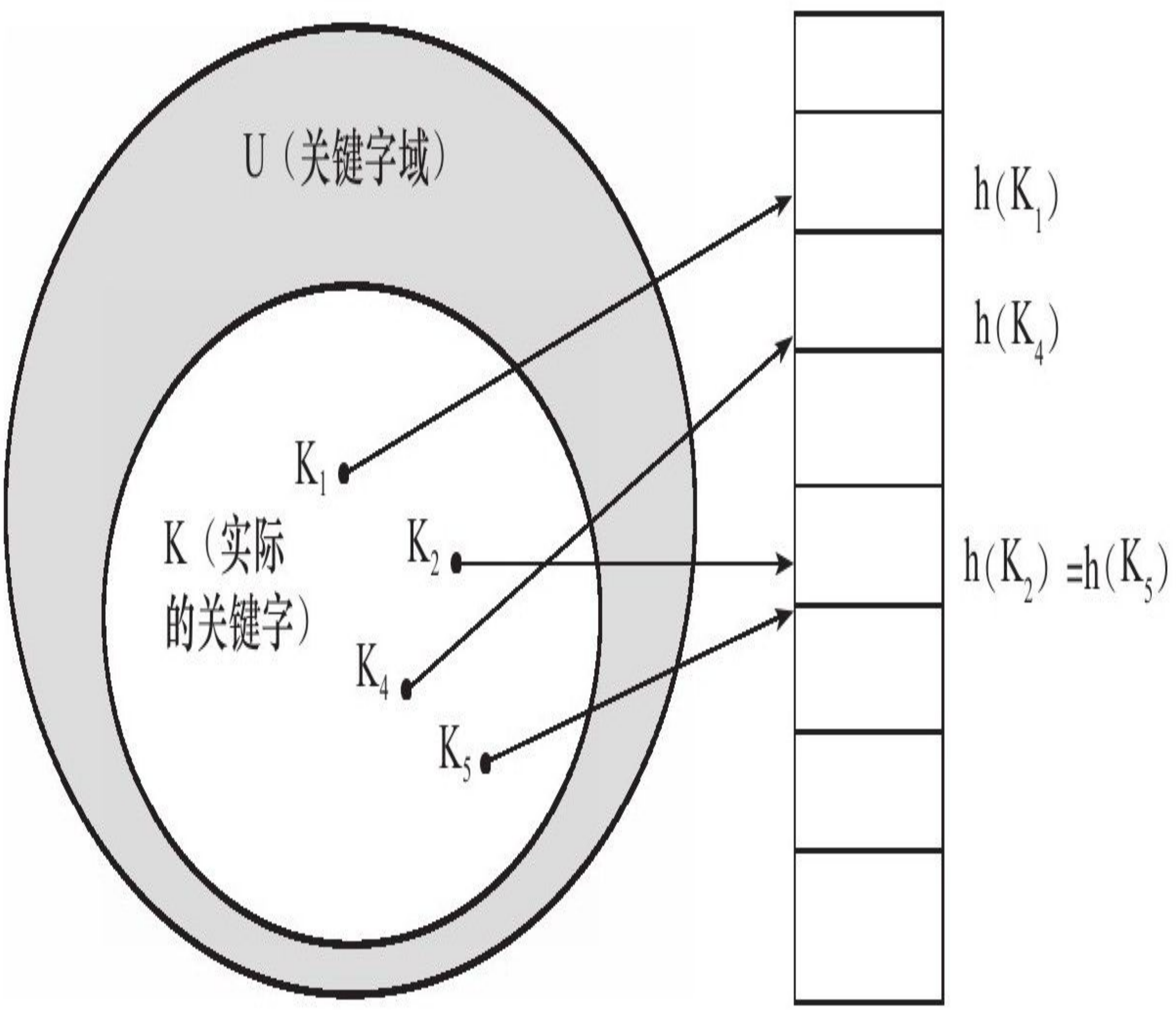
[illegible]T[0..m-1][illegible]U[illegible]5-38[illegible]k[illegible]k[illegible]k[illegible]T[k]=NULL[illegible]



[illegible] 5-38 [illegible]

[illegible]U[illegible]U[illegible]T[illegible]K[illegible]U[illegible]T[illegible]

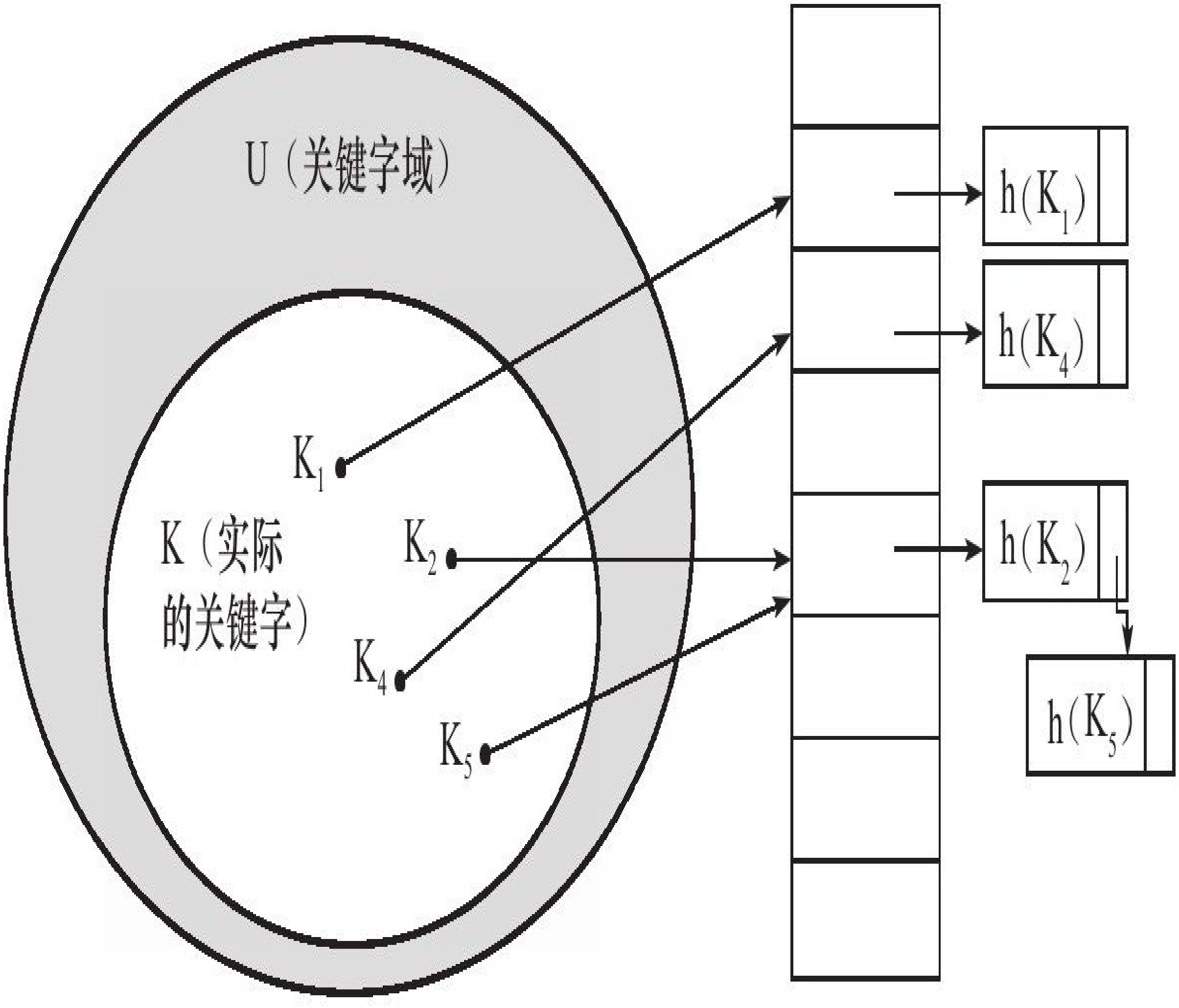
[illegible]h[illegible]k[illegible]h[illegible]k[illegible]h[illegible]U[illegible]T[0..m-1][illegible]5-39[illegible]



□ 5-39 □□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□5-39□□□collision□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□chaining□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□5-40□□□□j□□□□□□□□□□□□□□□□j□□□□□□□□□□□□□□□□□□□□□□□j□□NULL□



[illegible] 5-40 [illegible]

[illegible]h[illegible]

[illegible]k[illegible]m[illegible]k[illegible]m[illegible]

```
h(k)=k mod m
```

[1] [illegible]m[illegible]

5.7.2 InnoDB□□□□□□□□□□

InnoDB□□□Page□□□□□chain□□□□□□□□□□□□□□□□□□□□□□□□m□□□□□□□2□□□□□□□□□□□□□□□□□□□innodb_buffer_pool_size□□□□10M□□□□640□16KB□□□□□□□□□□□□□□□□□□□□□□640×2=1280□□□□□□□1280□□□□□□□□□1280□□□□□□□□□□□1399□□□□□□□□□□1399□□□□□□□□□□□□□□□□□□□□□□

□□InnoDB□□□

□□□□□□□InnoDB□□□□□□□□□□□□space_id□□□□□□□□□□□□□□□□□□□□□16KB□□□□□□□offset□InnoDB□□□□□space_id□□20□□□□□□□□space_id□offset□□□□□K=space_id□□20+space_id+offset□□□□□□□□□□□□□□□□

5.7.3 [illegible]

[illegible]DBA[illegible]SELECT*FROM TABLE WHERE index_col='xxx'[illegible]SHOW ENGINE INNODB STATUS[illegible]

```
mysql[illegible]SHOW ENGINE INNODB STATUS\G;

***************************1.row***************************

Status:

=====================================

090922 11:52:51 INNODB MONITOR OUTPUT

=====================================

Per second averages calculated from the last 15 seconds

……

-------------------------------------

INSERT BUFFER AND ADAPTIVE HASH INDEX

-------------------------------------

Ibuf:size 2249,free list len 3346,seg size 5596,

374650 inserts,51897 merged recs,14300 merges

Hash table size 4980499,node heap has 1246 buffer(s)

1640.60 hash searches/s,3709.46 non-hash searches/s

……
```

[illegible]

```
SELECT*FROM table WHERE index_col='xxx'
```

[illegible]non-hash searches/s[illegible]hash searches:non-hash searches[illegible]

[illegible]InnoDB[illegible]innodb_adaptive_hash_index[illegible]

5.8 [illegible]

5.8.1 [illegible]

[illegible]B+[illegible]prefix[illegible]B+[illegible]

```
SELECT*FROM blog WHERE content like'xxx%'
```

[illegible]SQL[illegible]xxx[illegible]content[illegible]B+[illegible]xxx[illegible]

```
SELECT*FROM blog WHERE content like'%xxx%'
```

[illegible]B+[illegible]SQL[illegible]B+[illegible]B+[illegible]

[illegible]Full-Text Search[illegible]

[illegible]MySQL[illegible]InnoDB[illegible]MyISAM[illegible]MyISAM[illegible]InnoDB[illegible]

[illegible]InnoDB 1.2.x[illegible]InnoDB[illegible]MyISAM[illegible]

5.8.2 □□□□

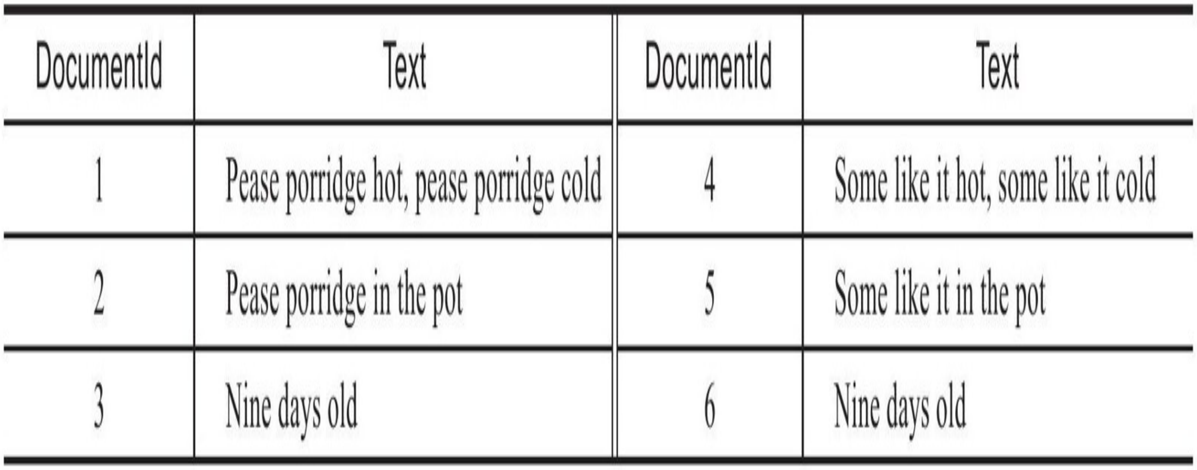
□□□□□□□□□□□□□inverted index□□□□□□□□□□B+□□□□□□□□□□□□□□□□□□□□□auxiliary table□□□

❑inverted file index□□□□□□□{□□□□□□□□□□ID}

❑full invertedindex□□□□□□□{□□□(□□□□□□□ID□□□□□□□□□□)}

□□□□□□□□□□□□□t□□□□□□□5-6□□□

表 5-6 全文检索表 t

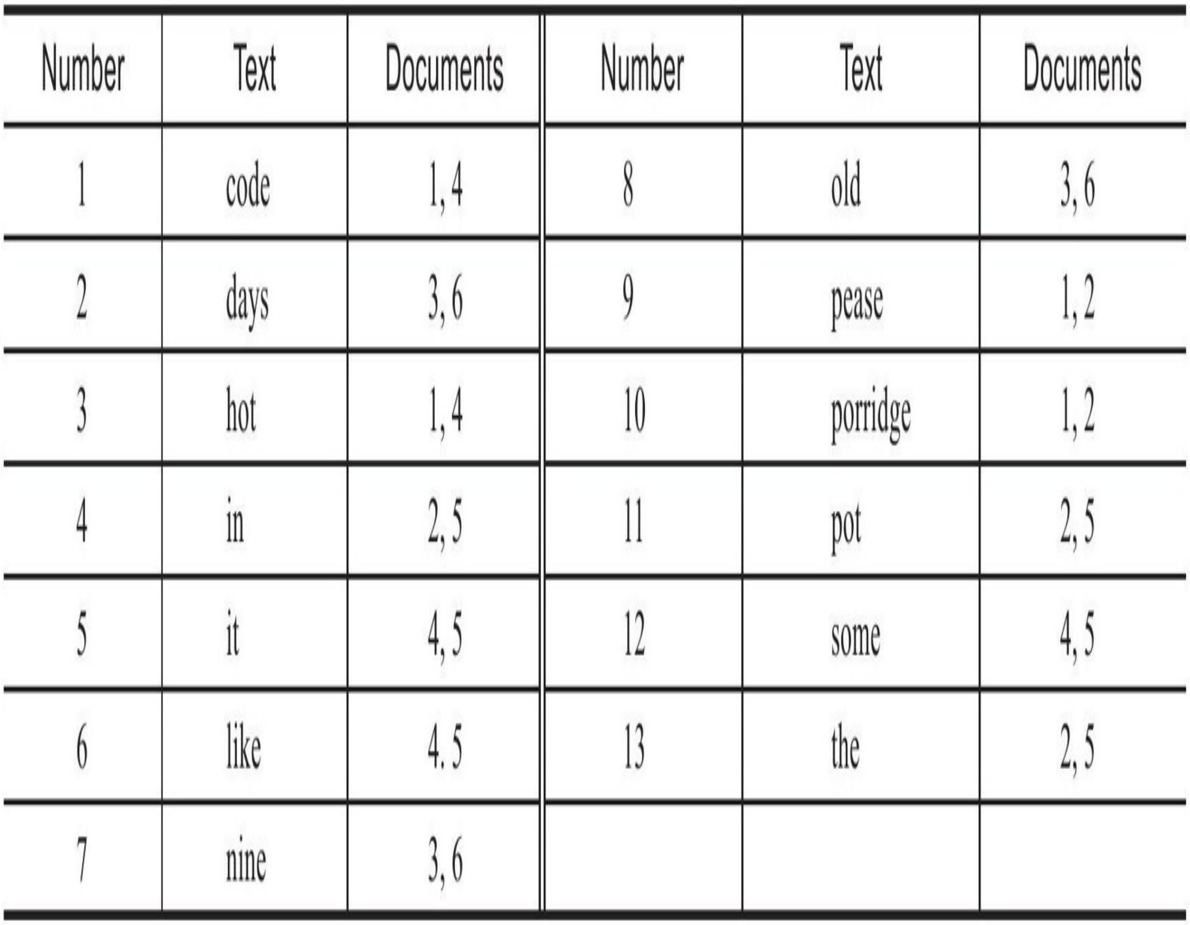


DocumentId	Text	DocumentId	Text
1	Pease porridge hot, pease porridge cold	4	Some like it hot, some like it cold
2	Pease porridge in the pot	5	Some like it in the pot
3	Nine days old	6	Nine days old

DocumentId□□□□□□□□□□□Id□Text□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Some□□□□□Id□□□□□□□□□□□□□□□□Some□□□□□Id□□□□

□□inverted file index□□□□□□□□□□□□□□5-7□□□

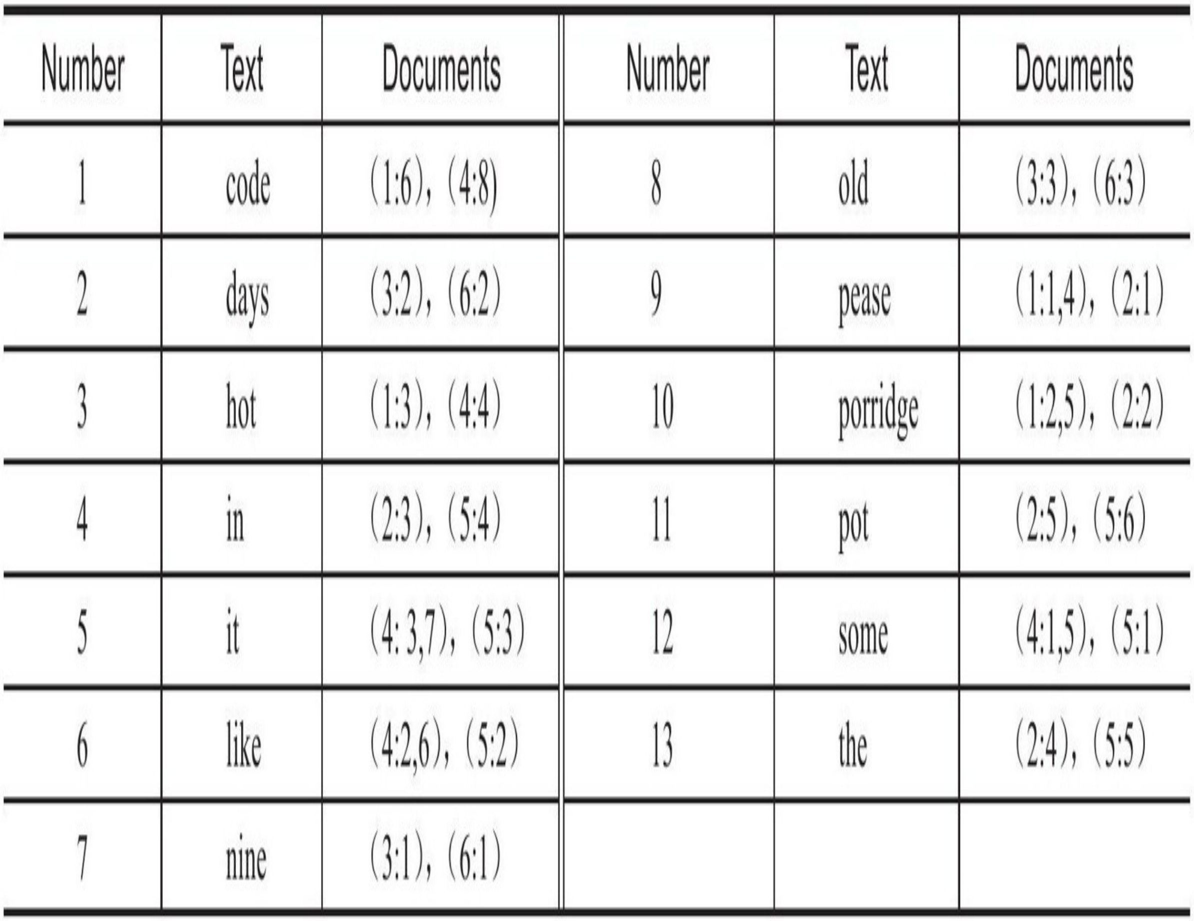
表 5-7 inverted file index 的关联数组



Number	Text	Documents	Number	Text	Documents
1	code	1, 4	8	old	3, 6
2	days	3, 6	9	pease	1, 2
3	hot	1, 4	10	porridge	1, 2
4	in	2, 5	11	pot	2, 5
5	it	4, 5	12	some	4, 5
6	like	4. 5	13	the	2, 5
7	nine	3, 6			

□□□□□code□□□□1□4□□□days□□□□3□6□□□□□□□□□□□□□□□□□□□Documents□□□□□□□□□□□□inverted file index□□□□□□Id□□full inverted index□□□□(pair)□□(DocumentId□Position)□□□□□□□□□□□5-8□□□

表 5-8 full inverted index 的关联数组



Number	Text	Documents	Number	Text	Documents
1	code	（1:6），（4:8）	8	old	（3:3），（6:3）
2	days	（3:2），（6:2）	9	pease	（1:1,4），（2:1）
3	hot	（1:3），（4:4）	10	porridge	（1:2,5），（2:2）
4	in	（2:3），（5:4）	11	pot	（2:5），（5:6）
5	it	（4: 3,7），（5:3）	12	some	（4:1,5），（5:1）
6	like	（4:2,6），（5:2）	13	the	（2:4），（5:5）
7	nine	（3:1），（6:1）			

full inverted index□□□□□□□□□□□□□□□code□□□□□□□□1:6□□□□□1□□6□□□□code□□□□□□full inverted index□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

5.8.3 InnoDB□□□□

InnoDB□□□□□1.2.x□□□□□□□□□□□□□□□□□full inverted index□□□□□InnoDB□□□□□□□(DocumentId□Position)□□□□“ilist”□□□□□□□□□□□□□□□□□□□□word□□□□□□□ilist□□□□□□word□□□□□□□□□□□□□□InnoDB□□□□□ilist□□□□□□Position□□□□□□□□Proximity Search□□MyISAM□□□□□□□□□□□

□□□□□□□□□□□□□□□□□word□□□□□□□□□□□□□Auxiliary Table□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□6□Auxiliary Table□□□□□□□□word□Latin□□□□□□□

Auxiliary Table□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□FTS Index Cache□□□□□□□□□□□□□□□□□□□□□□□□

FTS Index Cache□□□□□□□□□□□□□word□ilist□□□FTS Index Cache□□Auxiliary Table□□□□□□□□InnoDB□□□□□□□□Auxiliary Table□□□□□□□□□□□□□□□□□Auxiliary Table□□□□□□□□□□□□□Auxiliary Table□□□□□FTS Index Cache□□□□word□□□□□Auxiliary Table□□□□□□□□□□□□merge□□□□□□□□□□□Insert Buffer□□□□□□□□Insert Buffer□□□□□□□□□□□□□B+□□□□□□□FTS Index Cache□□□□□Insert Buffer□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Auxiliary Table□□□□□

InnoDB□□□□□□□□□□□□□□□□□Auxiliary Table□□□□□□□□□□□□□□□innodb_ft_aux_table□□□□□□□□Auxiliary Table□□□□SQL□□□□□□test□□□□fts_a□Auxiliary Table□

```
mysql□SET GLOBAL innodb_ft_aux_table='test/fts_a';

Query OK,0 rows affected(0.00 sec)
```

□□□□□□□□□□□□□□□□information_schema□□□□□INNODB_FT_INDEX_TABLE□□□fts_a□□□□□□□

[illegible]Oracle 11g[illegible]InnoDB[illegible]FTS Index Cache[illegible]InnoDB[illegible]

[illegible]FTS Index Cache[illegible]Auxiliary Table[illegible]FTS Index Cache[illegible]InnoDB[illegible]FTS Index Cache[illegible]

[illegible]innodb_ft_cache_size[illegible]FTS Index Cache[illegible]32M[illegible](word[illegible]ilist)[illegible]Auxiliary Table[illegible]

FTS Document ID[illegible]InnoDB[illegible]word[illegible]InnoDB[illegible]FTS_DOC_ID[illegible]BIGINT UNSIGNED NOT NULL[illegible]InnoDB[illegible]FTS_DOC_ID_INDEX[illegible]Unique Index[illegible]InnoDB[illegible]FTS_DOC_ID[illegible]Unique Index[illegible]FTS_DOC_ID[illegible]MySQL[illegible]

```
mysql[illegible]CREATE TABLE fts_a(

-[illegible]FTS_DOC_ID INT UNSIGNED AUTO_INCREMENT NOT NULL,

-[illegible]body TEXT,

-[illegible]PRIMARY KEY(FTS_DOC_ID)

-[illegible]);

ERROR 1166(42000):Incorrect column name'FTS_DOC_ID'
```

[illegible]FTS_DOC_ID[illegible]INT[illegible]BIG INT[illegible]Incorrect column name'FTS_DOC_ID'[illegible]

```
mysql[illegible]CREATE TABLE fts_a(
-□FTS_DOC_ID BIGINT UNSIGNED AUTO_INCREMENT NOT NULL,

-□body TEXT,

-□PRIMARY KEY(FTS_DOC_ID)

-□);

Query OK,0 rows affected(0.02 sec)
```

□□Auxiliary Table□□□□□□□□□□FTS Cache Index□□□□□□□Auxiliary Table□□□□□□□□InnoDB□□□□□□□□FTS Document ID□□□□□□□DELETED auxiliary table□□□□□□□innodb_ft_aux_table□□□□□□□□□□information_schema□□□□□INNODB_FT_DELETED□□□□□□FTS Document ID□

□□□□□DML□□□□□□□□□□□□□□□□□□□□□□□DELETED□□InnoDB□□□OPTIMIZE TABLE□□□OPTIMIZE TABLE□□□□□□□□□□□□□Cardinality□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□innodb_optimize_fulltext_only□□□□□□□

```
mysql□SET GLOBAL innodb_optimize_fulltext_only=1;

mysql□OPTIMIZE TABLEfts_a;
```

□□□□□□□□□□□□□OPTIMIZE TABLE□□innodb_ft_num_word_optimize□□□□□□□□□□□□□□□□□□□□□□□2000□

□□□□□□□□□□□□□□□□□□□□□□□fts_a□

```
CREATE TABLE fts_a(

FTS_DOC_ID BIGINT UNSIGNED AUTO_INCREMENT NOT NULL,

body TEXT,
PRIMARY KEY(FTS_DOC_ID)

);

INSERT INTO fts_a

SELECT NULL,'Pease porridge in the pot';

INSERT INTO fts_a

SELECT NULL,'Pease porridge hot,pease porridge cold';

INSERT INTO fts_a

SELECT NULL,'Nine days old';

INSERT INTO fts_a

SELECT NULL,'Some like it hot,some like it cold';

INSERT INTO fts_a

SELECT NULL,'Some like it in the pot';

INSERT INTO fts_a

SELECT NULL,'Nine days old';

INSERT INTO fts_a

SELECT NULL,'I like code days';

CREATE FULLTEXT INDEX idx_fts ON fts_a(body);
```

[illegible]fts_a[illegible]body[illegible]FULLTEXT[illegible] fts_a[illegible]

```
mysql[illegible]SELECT*FROM fts_a;

+------------+---------------------------------------+

|FTS_DOC_ID|body|

+------------+---------------------------------------+

|1|Pease porridge in the pot|

|2|Pease porridge hot,pease porridge cold|

|3|Nine days old|

|4|Some like it hot,some like it cold|
|5|Some like it in the pot|

|6|Nine days old|

|7|I like code days|

+------------+----------------------------------------+

7 rows in set(0.00 sec)
```

□□□□□□innodb_ft_aux_table□□□□□□□□□□□

```
mysql□SET GLOBAL innodb_ft_aux_table='test/fts_a';

Query OK,0 rows affected(0.00 sec)

mysql□SELECT*FROM information_schema.INNODB_FT_INDEX_TABLE;

+----------+--------------+-------------+-----------+--------+----------+

|WORD|FIRST_DOC_ID|LAST_DOC_ID|DOC_COUNT|DOC_ID|POSITION|

+----------+--------------+-------------+-----------+--------+----------+

|code|7|7|1|7|7|

|cold|2|4|2|2|35|

|cold|2|4|2|4|31|

|days|3|7|3|3|5|

|days|3|7|3|6|5|

|days|3|7|3|7|12|

|hot|2|4|2|2|15|

|hot|2|4|2|4|13|

|like|4|7|3|4|5|

|like|4|7|3|4|18|

|like|4|7|3|5|5|

|like|4|7|3|7|2|

|nine|3|6|2|3|0|

|nine|3|6|2|6|0|

|old|3|6|2|3|10|

|old|3|6|2|6|10|
|pease|1|2|2|1|0|

|pease|1|2|2|2|0|

|pease|1|2|2|2|20|

|porridge|1|2|2|1|6|

|porridge|1|2|2|2|6|

|porridge|1|2|2|2|20|

|pot|1|5|2|1|22|

|pot|1|5|2|5|20|

|some|4|5|2|4|0|

|some|4|5|2|4|18|

|some|4|5|2|5|0|

+----------+--------------+------------+-----------+--------+----------+

27 rows in set(0.00 sec)
```

□□□□□□word□□□□□□DOC_ID□POSITION□□□□□□□□FIRST_DOC_ID□LAST_DOC_ID□□DOC_COUNT□□□□□□□word□□□□□□□□ID□□□□□□□□□□ID□□□□word□□□□□□□□□□

□□□□□□□□SQL□□□□□□FTS_DOC_ID□7□□□□

```
mysql□DELETE FROM test.fts_a WHERE FTS_DOC_ID=7;

Query OK,1 row affected(0.00 sec)
```

□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□ID□□□DELETED□□□□□□□□□□□□□□□□

```
mysql□SELECT*FROM INNODB_FT_DELETED;

+--------+

|DOC_ID|

+--------+

|7|
+--------+

1 row in set(0.00 sec)
```

□□□□□□□□□ID□□□□□INNODB_FT_DELETED□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□SQL□□□

```
mysql□SET GLOBAL innodb_optimize_fulltext_only=1;

Query OK,0 rows affected(0.00 sec)

mysql□OPTIMIZE TABLE test.fts_a;

+------------+----------+----------+----------+

|Table|Op|Msg_type|Msg_text|

+------------+----------+----------+----------+

|test.fts_a|optimize|status|OK|

+------------+----------+----------+----------+

1 row in set(0.01 sec)

mysql□SELECT*FROM INNODB_FT_DELETED;

+--------+

|DOC_ID|

+--------+

|7|

+--------+

1 row in set(0.00 sec)

mysql□SELECT*FROM INNODB_FT_BEING_DELETED;

+--------+

|DOC_ID|

+--------+

|7|

+--------+

1 row in set(0.00 sec)
```

[illegible]OPTIMIZE TABLE[illegible]ID[illegible]INNODB_FT_BEING_DELETED[illegible]7[illegible]ID[illegible]ID[illegible]

```
mysql[illegible]INSERT INTO test.fts_a SELECT 7,'I like this days';

ERROR 182(HY000):Invalid InnoDB FTS Doc ID
```

stopword[illegible]stopword list[illegible]word[illegible]the[illegible]stopword[illegible]InnoDB[illegible]stopword[illegible]information_schema[illegible]INNODB_FT_DEFAULT_STOPWORD[illegible]36[illegible]stopword[illegible]innodb_ft_server_stopword_table[illegible]stopword[illegible]

```
mysql[illegible]CREATE TABLE user_stopword(

-[illegible]value VARCHAR(30)

-[illegible])ENGINE=INNODB;

Query OK,0 rows affected(0.03 sec)

mysql[illegible]SET GLOBAL

-[illegible]innodb_ft_server_stopword_table="test/user_stopword";

Query OK,0 rows affected(0.00 sec)
```

[illegible]InnoDB[illegible]

❑[illegible]

❑[illegible]

❑[illegible]delimiter[illegible]

5.8.4 □□□□

MySQL□□□□□□□□□□Full-Text Search□□□□□□□□□□

```
MATCH(col1,col2,...)AGAINST(expr[search_modifier])

search_modifier:

{

IN NATURAL LANGUAGE MODE

|IN NATURAL LANGUAGE MODE WITH QUERY EXPANSION

|IN BOOLEAN MODE

|WITH QUERY EXPANSION

}
```

MySQL□□□□□MATCH()…AGAINST()□□□□□□□□□□□□MATCH□□□□□□□□□□□AGAINST□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

1.Natural Language

□□□□□□MATCH□□□□□□□□□□□Natural Language□□□□□□□□□□□□word□□□□□□5.8.3□□□□□□□fts_a□□□body□□□□□Pease□□□□□□□□□□□□□□□□□□□□□□SQL□□□

```
mysql□SELECT*FROM fts_a WHERE body LIKE'%Pease%';
```

□□□□SQL□□□□□□B+□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□

```
mysql□SELECT*FROM fts_a

-□WHERE MATCH(body)

-□AGAINST('Porridge'IN NATURAL LANGUAGE MODE);

+------------+----------------------------------------+
|FTS_DOC_ID|body|

+------------+----------------------------------------+

|2|Pease porridge hot,pease porridge cold|

|1|Pease porridge in the pot|

+------------+----------------------------------------+

2 rows in set(0.00 sec)
```

[illegible]NATURAL LANGUAGE MODE[illegible]SQL[illegible]

```
SELECT*FROM fts_a WHERE MATCH(body)AGAINST('Porridge');
```

[illegible]SQL[illegible]

```
mysql[illegible]EXPLAIN SELECT*FROM fts_a

-[illegible]WHERE MATCH(body)AGAINST('Porridge')\G;

***************************1.row***************************

id:1

select_type:SIMPLE

table:fts_a

type:fulltext

possible_keys:idx_fts

key:idx_fts

key_len:0

ref:NULL

rows:1

Extra:Using where

1 row in set(0.00 sec)
```

[illegible]type[illegible]fulltext[illegible]key[illegible]idx_fts[illegible]MATCH[illegible]

```
mysql[illegible]SELECT*FROM fts_b
-[illegible]WHERE MATCH(body)AGAINST('Porridge');
ERROR 1191(HY000):Can't find FULLTEXT index matching the column list
```

[illegible]WHERE[illegible]MATCH[illegible]Relevance[illegible]0[illegible]MySQL[illegible]

❑word[illegible]

❑word[illegible]

❑word[illegible]

❑[illegible]word[illegible]

[illegible]Porridge[illegible]2[illegible]

[illegible]MATCH[illegible]SQL[illegible]

```
mysql[illegible]SELECT count(*)
-[illegible]FROM fts_a WHERE
-[illegible]MATCH(body)AGAINST('Porridge'IN NATURAL LANGUAGE MODE);
+-------------------+
|count(FTS_DOC_ID)|
+-------------------+
|2|
+-------------------+
1 row in set(0.00 sec)
```

□□SQL□□□□□□□□□

```
mysql□SELECT
-□COUNT(IF(MATCH(body)
-□AGAINST('Porridge'IN NATURAL LANGUAGE MODE),1,NULL))
-□AS count
-□FROM fts_a;
+-------+
|count|
+-------+
|2|
+-------+
1 row in set(0.00 sec)
```

□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□

□□□□□□□□□SQL□□□□□□□□

```
mysql□SELECT fts_doc_id,body,
-□MATCH(body)AGAINST('Porridge'IN NATURAL LANGUAGE MODE)
-□AS Relevance
-□FROM fts_a;
+------------+----------------------------------------+-------------------+
|fts_doc_id|body|Relevance|
+------------+----------------------------------------+-------------------+
|1|Pease porridge in the pot|0.2960100471973419|
|2|Pease porridge hot,pease porridge cold|0.5920200943946838|
|3|Nine days old|0|
|4|Some like it hot,some like it cold|0|

|5|Some like it in the pot|0|

|6|Nine days old|0|

|7|I like hot and code days|0|

+------------+---------------------------------------+-------------------+

7 rows in set(0.01 sec)
```

[illegible]InnoDB[illegible]

❑[illegible]word[illegible]stopword[illegible]

❑[illegible]word[illegible][innodb_ft_min_token_size[illegible]innodb_ft_max_token_size][illegible]

[illegible]stopword[illegible]the[illegible]

```
mysql[illegible]SELECT fts_doc_id AS id,body,

-[illegible]MATCH(body)AGAINST('the'IN NATURAL LANGUAGE MODE)

-[illegible]AS rl

-[illegible]FROM fts_a;

+----+---------------------------------------+------+

|id|body|rl|

+----+---------------------------------------+------+

|1|Pease porridge in the pot|0|

|2|Pease porridge hot,pease porridge cold|0|

|3|Nine days old|0|

|4|Some like it hot,some like it cold|0|

|5|Some like it in the pot|0|

|6|Nine days old|0|

|7|I like hot and code days|0|

+----+---------------------------------------+------+
7 rows in set(0.00 sec)
```

□□□□□the□□□□□1□5□□□□□□□□□stopword□□□□□□□0□

□□innodb_ft_min_token_size□innodb_ft_max_token_size□□InnoDB□□□□□□□□□□□□□□□□□innodb_ft_min_token_size□□□□□□□innodb_ft_max_token_size□□□□□□□□□□□□InnoDB□□□□□□□□innodb_ft_min_token_size□□□□□3□□□innodb_ft_max_token_size□□□□□84□

2.Boolean

MySQL□□□□□□□IN BOOLEAN MODE□□Pease□□□hot□□□□□□□+□-□□□□□□□□□□□□□□□□□□□□□

```
mysql□SELECT*FROM fts_a

-□WHERE MATCH(body)AGAINST('+Pease-hot'IN BOOLEAN MODE)\G;

***************************1.row***************************

FTS_DOC_ID:1

body:Pease porridge in the pot
```

Boolean□□□□□□□□□□□□□□

❑+□□□word□□□□□

❑-□□□word□□□□□□

❑□no operator□□□□word□□□□□□□□□□□□□□□□□□□

❑@distance□□□□□□□□□□□□□□□□□distance□□□distance□□□□□□□□□□□□□□□□□□□Proximity Search□□MATCH□body□AGAINST□'"Pease pot"@30'IN BOOLEAN MODE□□□□□□Pease□pot□□□□□□□30□□□□

❑□□□□□□□□□□□□□□□

❑□□□□□□□□□□□□□□□

❑□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

❑*□□□□□□□□□□□□□lik*□□□□□□lik□like□□□□likes□

❑"□□□□□

□□□□□□□□□□□□□□□□□□fts_a□□□□□□□□□□□□SQL□□□□□pease□□hot□□□□

```
mysql□SELECT*FROM fts_a

-□WHERE MATCH(body)AGAINST('+Pease+hot'IN BOOLEAN MODE)\G;

***************************1.row***************************

FTS_DOC_ID:2

body:Pease porridge hot,pease porridge cold

1 row in set(0.00 sec)
```

□□□SQL□□□□□pease□□□hot□□□□

```
mysql□SELECT*FROM fts_a

-□WHERE MATCH(body)AGAINST('+Pease-hot'IN BOOLEAN MODE)\G;

***************************1.row***************************

FTS_DOC_ID:1

body:Pease porridge in the pot

1 row in set(0.00 sec)
```

□□□SQL□□□□□pease□□hot□□□□

```
mysql□SELECT*FROM fts_a
```

```
-□WHERE MATCH(body)AGAINST('Pease hot'IN BOOLEAN MODE);

+------------+----------------------------------------+

|FTS_DOC_ID|body|

+------------+----------------------------------------+

|2|Pease porridge hot,pease porridge cold|

|1|Pease porridge in the pot|

|4|Some like it hot,some like it cold|

|7|I like hot and code days|

+------------+----------------------------------------+

4 rows in set(0.00 sec)
```

□□□SQL□□□□Proximity Search□

```
mysql□SELECT fts_doc_id,body FROM fts_a

-□WHERE MATCH(body)

-□AGAINST('"Pease pot"@30'IN BOOLEAN MODE)\G;

***************************1.row***************************

fts_doc_id:1

body:Pease porridge in the pot

1 row in set(0.01 sec)

mysql□SELECT fts_doc_id,body FROM fts_a

-□WHERE MATCH(body)

-□AGAINST('"Pease pot"@10'IN BOOLEAN MODE);

Empty set(0.01 sec)
```

□□□□□□1□□□Pease□pot□□□□22□□□□□□□□@30□□□□□□□□□□□□□@10□□□□□□□□□□□□□□

```
mysql□SELECT fts_doc_id,body,

-□MATCH(body)AGAINST('like□pot'IN BOOLEAN MODE)
```

```
-□AS Relevance FROM fts_a;

+------------+----------------------------------------+--------------------+

|fts_doc_id|body|Relevance|

+------------+----------------------------------------+--------------------+

|1|Pease porridge in the pot|1.2960100173950195|

|2|Pease porridge hot,pease porridge cold|0|

|3|Nine days old|0|

|4|Some like it hot,some like it cold|0.27081382274627686|

|5|Some like it in the pot|1.4314169883728027|

|6|Nine days old|0|

|7|I like hot and code days|0.13540691137313843|

+------------+----------------------------------------+--------------------+

7 rows in set(0.00 sec)
```

□□SQL□□□□□□□□□□□like□pot□□□□□□□□□□□□□□pot□□□□□□□□□□□4□□□□□□like□□□□□□□pot□□□□□□□□□□1□5□□

□□□□□□□□“□some”□□□□□□□□□□□□

```
mysql□SELECT fts_doc_id,body,

-□MATCH(body)AGAINST('like□hot□some'IN BOOLEAN MODE)

-□AS Relevance

-□FROM fts_a;

+------------+----------------------------------------+--------------------+

|fts_doc_id|body|Relevance|

+------------+----------------------------------------+--------------------+

|1|Pease porridge in the pot|0|

|2|Pease porridge hot,pease porridge cold|1.2960100173950195|

|3|Nine days old|0|

|4|Some like it hot,some like it cold|1.158843994140625|
|5|Some like it in the pot|-0.5685830116271973|

|6|Nine days old|0|

|7|I like hot and code days|0.13540691137313843|

+------------+---------------------------------------+--------------------+

7 rows in set(0.00 sec)
```

□□□□□□5□□□□□□□□□□□□□□□□□□□like□□□□□□□□some□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□SQL□□□

```
mysql□SELECT*FROM fts_a

-□WHERE MATCH(body)AGAINST('po*'IN BOOLEAN MODE);

+------------+---------------------------------------+

|FTS_DOC_ID|body|

+------------+---------------------------------------+

|2|Pease porridge hot,pease porridge cold|

|1|Pease porridge in the pot|

|5|Some like it in the pot|

+------------+---------------------------------------+

3 rows in set(0.00 sec)
```

□□□□□□□□□□□□□□□po□□□□□□□porridge□pot□

□□□□□□□□SQL□□□□□

```
mysql□SELECT*FROM fts_a

-□WHERE MATCH(body)AGAINST('like hot'IN BOOLEAN MODE);

+------------+---------------------------------------+

|FTS_DOC_ID|body|

+------------+---------------------------------------+
|4|Some like it hot,some like it cold|

|7|I like hot and code days|

|2|Pease porridge hot,pease porridge cold|

|5|Some like it in the pot|

+------------+---------------------------------------+

4 rows in set(0.00 sec)

mysql□SELECT*FROM fts_a

-□WHERE MATCH(body)AGAINST('"like hot"'IN BOOLEAN MODE);

+------------+--------------------------+

|FTS_DOC_ID|body|

+------------+--------------------------+

|7|I like hot and code days|

+------------+--------------------------+

1 row in set(0.00 sec)
```

□□□□□□□SQL□□□□□□""□like□hot□□□□□□□□□□□□□□□□□□□□□□□□□□□4□□□□□□□□SQL□□□"like hot"□□□□□□□□□□□□□□4□□□□□□□

3.Query Expansion

MySQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□implied knowledge□□□□□□□□□□□□□□□□□database□□□□□□□□□□□□□□□□□□□database□□□□□□□□□□□□MySQL□Oracle□DB2□RDBMS□□□□□□□□□□□Query Expansion□□□□□□□□□□implied knowledge□

□□□□□□□□□□WITH QUERY EXPANSION□IN NATURAL LANGUAGE MODE WITH QUERY EXPANSION□□□□blind query expansion□□□□automatic relevance feedback□□□□□□□□□□□□

❑□□□□□□□□□□□□□□□□□□□□□

❑□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□articles□

```
CREATE TABLE articles(

id INT UNSIGNED AUTO_INCREMENT NOT NULL PRIMARY KEY,

title VARCHAR(200),

body TEXT,

FULLTEXT(title,body)

)ENGINE=InnoDB;

INSERT INTO articles(title,body)VALUES

('MySQL Tutorial','DBMS stands for DataBase...'),

('How To Use MySQL Well','After you went through a...'),

('Optimizing MySQL','In this tutorial we will show...'),

('1001 MySQL Tricks','1.Never run mysqld as root.2....'),

('MySQL vs.YourSQL','In the following database comparison...'),

('MySQL Security','When configured properly,MySQL...'),

('Tuning DB2','For IBM database...'),

('IBM History','DB2 hitory for IBM...');
```

□□□□□□□□□□□□□□FTS_DOC_ID□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□articles□□□□□□□□□□□title□body□□□□□□□□□□database□□□□□□□□□□□□□

```
mysql□SELECT*FROM articles

-□WHERE MATCH(title,body)

-□AGAINST('database'IN NATURAL LANGUAGE MODE);

+----+-------------------+------------------------------------------+

|id|title|body|

+----+-------------------+------------------------------------------+

|1|MySQL Tutorial|DBMS stands for DataBase...|

|5|MySQL vs.YourSQL|In the following database comparison...|

|7|Tuning DB2|For IBM database...|
+----+------------------+---------------------------------------+

3 rows in set(0.00 sec)
```

□□□□□□□□□□3□□□□body□□□□database□□□□□□□□Query Expansion□□□□□□□□□□□□□□□

```
mysql□SELECT*FROM articles

-□WHERE MATCH(title,body)

-□AGAINST('database'WITH QUERY EXPANSION);

+----+----------------------+---------------------------------------+

|id|title|body|

+----+----------------------+---------------------------------------+

|5|MySQL vs.YourSQL|In the following database comparison...|

|1|MySQL Tutorial|DBMS stands for DataBase...|

|7|Tuning DB2|For IBM database...|

|8|IBM History|DB2 hitory for IBM...|

|3|Optimizing MySQL|In this tutorial we will show...|

|6|MySQL Security|When configured properly,MySQL...|

|2|How To Use MySQL Well|After you went through a...|

|4|1001 MySQL Tricks|1.Never run mysqld as root.2....|

+----+----------------------+---------------------------------------+

8 rows in set(0.00 sec)
```

□□□□□□□□8□□□□□□□□□□database□□□□□□□□title□body□□□□□ MySQL□DB2□□□□□□□Query Expansion□

□□Query Expansion□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

5.9 [illegible]

[illegible]B+[illegible]InnoDB1.2[illegible]B+[illegible]

□6□　□

□□□locking□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□MySQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Oracle□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□

□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□MyISAM□Oracle□SQL Server□□□□□□□□□□□□□□□□□□□□□□□□□□□“□□”□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

6.1 [illegible]

[illegible][1][illegible]InnoDB[illegible]InnoDB[illegible]LRU[illegible]LRU[illegible]

[illegible]SQL[illegible]SQL[illegible]InnoDB[illegible]MySQL[illegible]MyISAM[illegible]NDB Cluster[illegible]MySQL[illegible]Microsoft SQL Server[illegible]Oracle[illegible]

[illegible]MyISAM[illegible]“[illegible]”[illegible]MyISAM[illegible]Microsoft SQL Server[illegible]Microsoft SQL Server 2005[illegible]MyISAM[illegible]2005[illegible]Microsoft SQL Server[illegible]InnoDB[illegible]Microsoft SQL Server[illegible]

InnoDB[illegible]Oracle[illegible]

[1] [illegible]“[illegible]”[illegible]“[illegible]”[illegible]

6.2 lock[illegible]latch

[illegible]lock[illegible]latch[illegible]lock[illegible]latch[illegible]“[illegible]”[illegible]lock[illegible]

latch[illegible]InnoDB[illegible]latch[illegible]mutex[illegible]rwlock[illegible]

lock[illegible]lock[illegible]commit[illegible]rollback[illegible]lock[illegible]6-1[illegible]lock[illegible]latch[illegible]

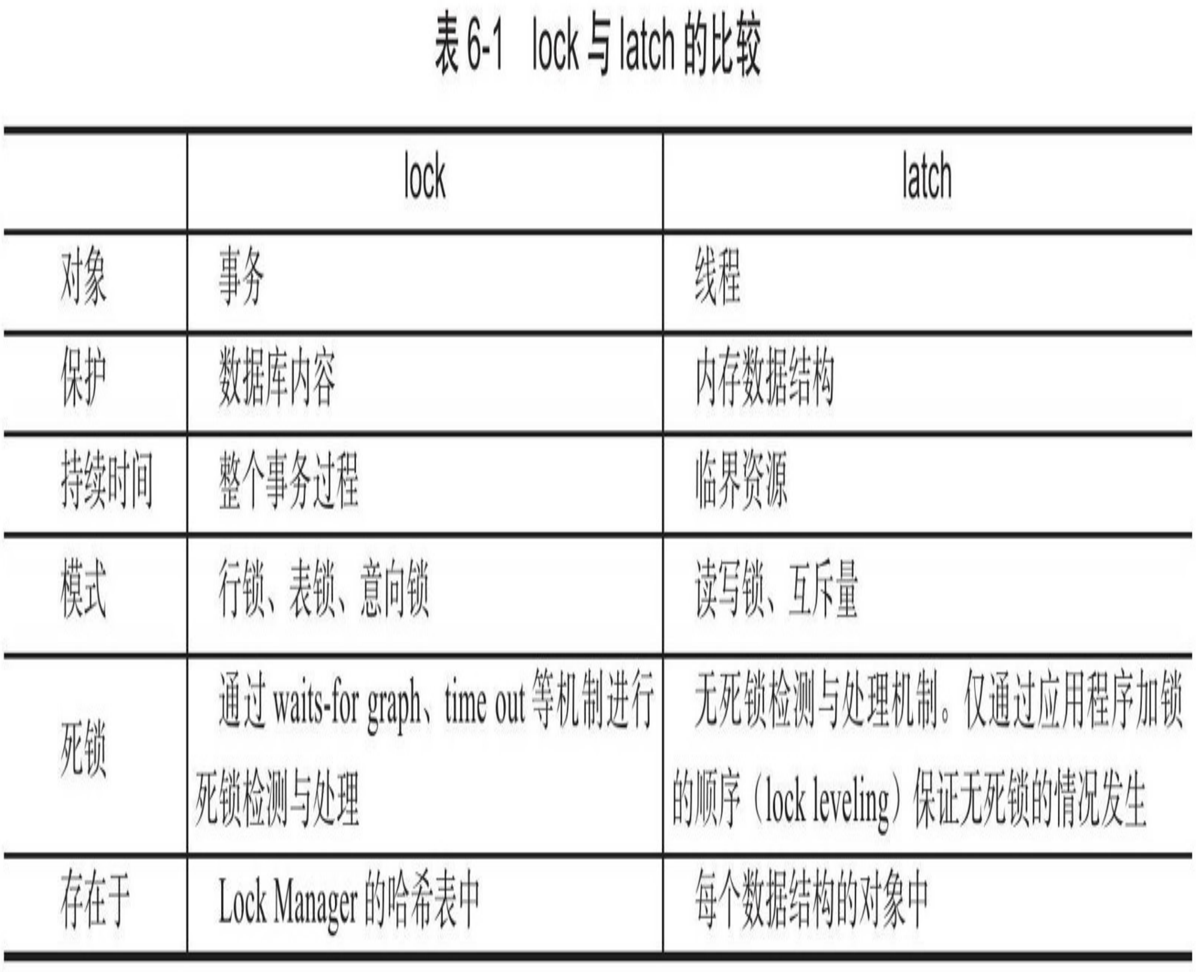
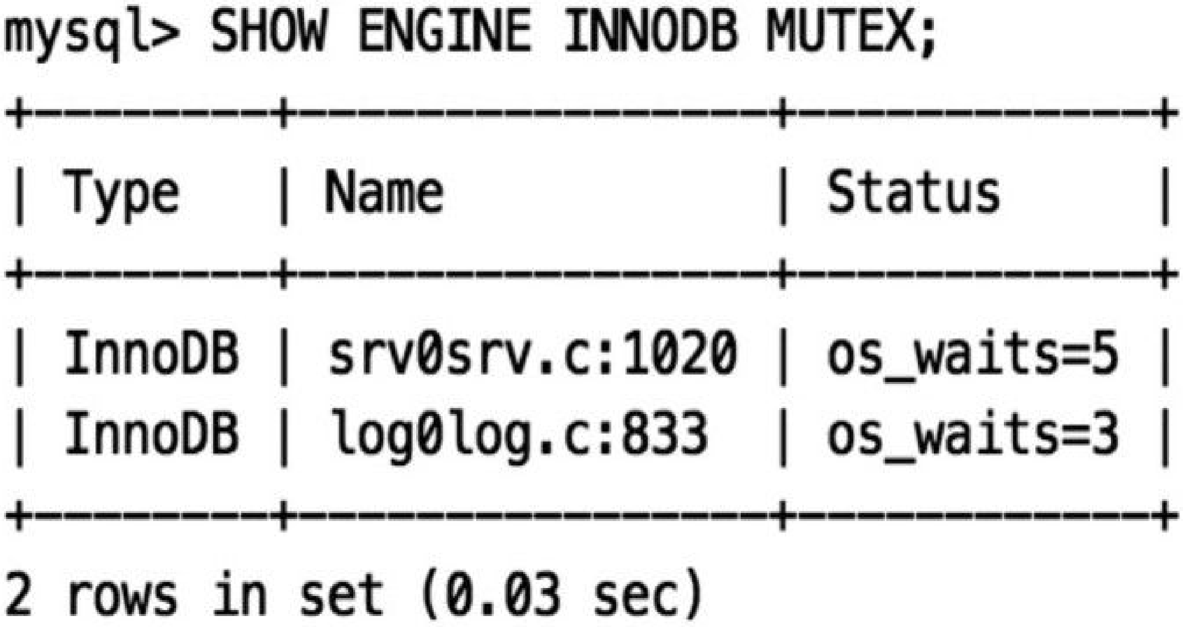


表6-1 lock与latch的比较

	lock	latch
对象	事务	线程
保护	数据库内容	内存数据结构
持续时间	整个事务过程	临界资源
模式	行锁、表锁、意向锁	读写锁、互斥量
死锁	通过 waits-for graph、time out 等机制进行死锁检测与处理	无死锁检测与处理机制。仅通过应用程序加锁的顺序（lock leveling）保证无死锁的情况发生
存在于	Lock Manager 的哈希表中	每个数据结构的对象中

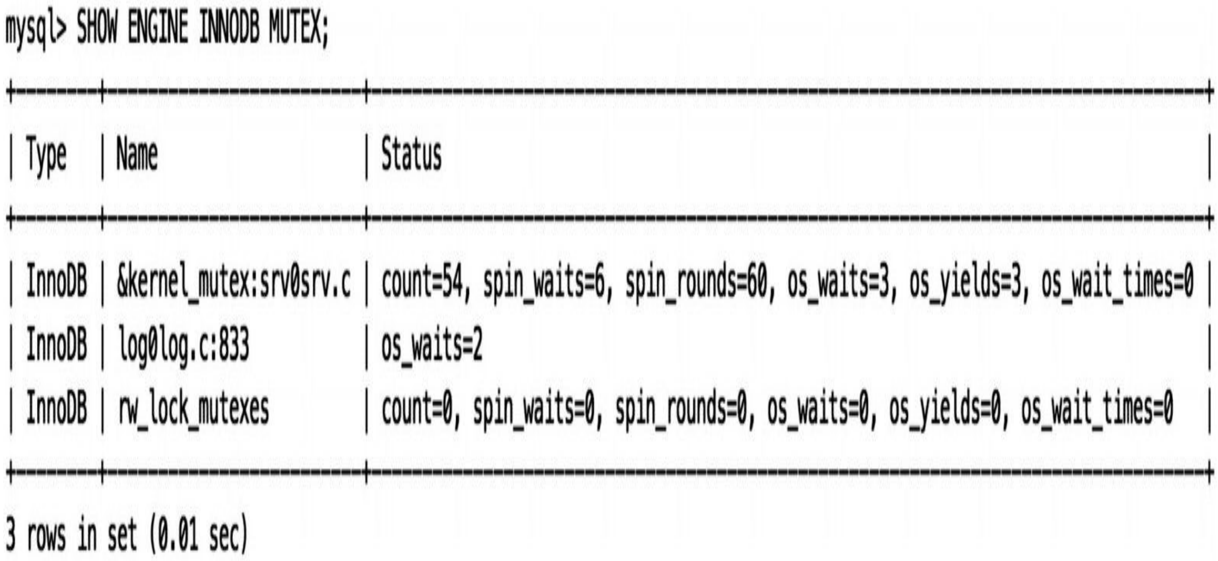
□□InnoDB□□□□□□latch□□□□□□□SHOW ENGINE INNODB MUTEX □□□□□□□□6-1□□□



```
mysql> SHOW ENGINE INNODB MUTEX;
+--------+-----------------+------------+
| Type   | Name            | Status     |
+--------+-----------------+------------+
| InnoDB | srv0srv.c:1020  | os_waits=5 |
| InnoDB | log0log.c:833   | os_waits=3 |
+--------+-----------------+------------+
2 rows in set (0.03 sec)
```

□ 6-1 □□□□SHOW ENGINE INNODB MUTEX□□latch

□Debug□□□□□□□□SHOW ENGINE INNODB MUTEX□□□□latch□□□□□□□□6-2□□□

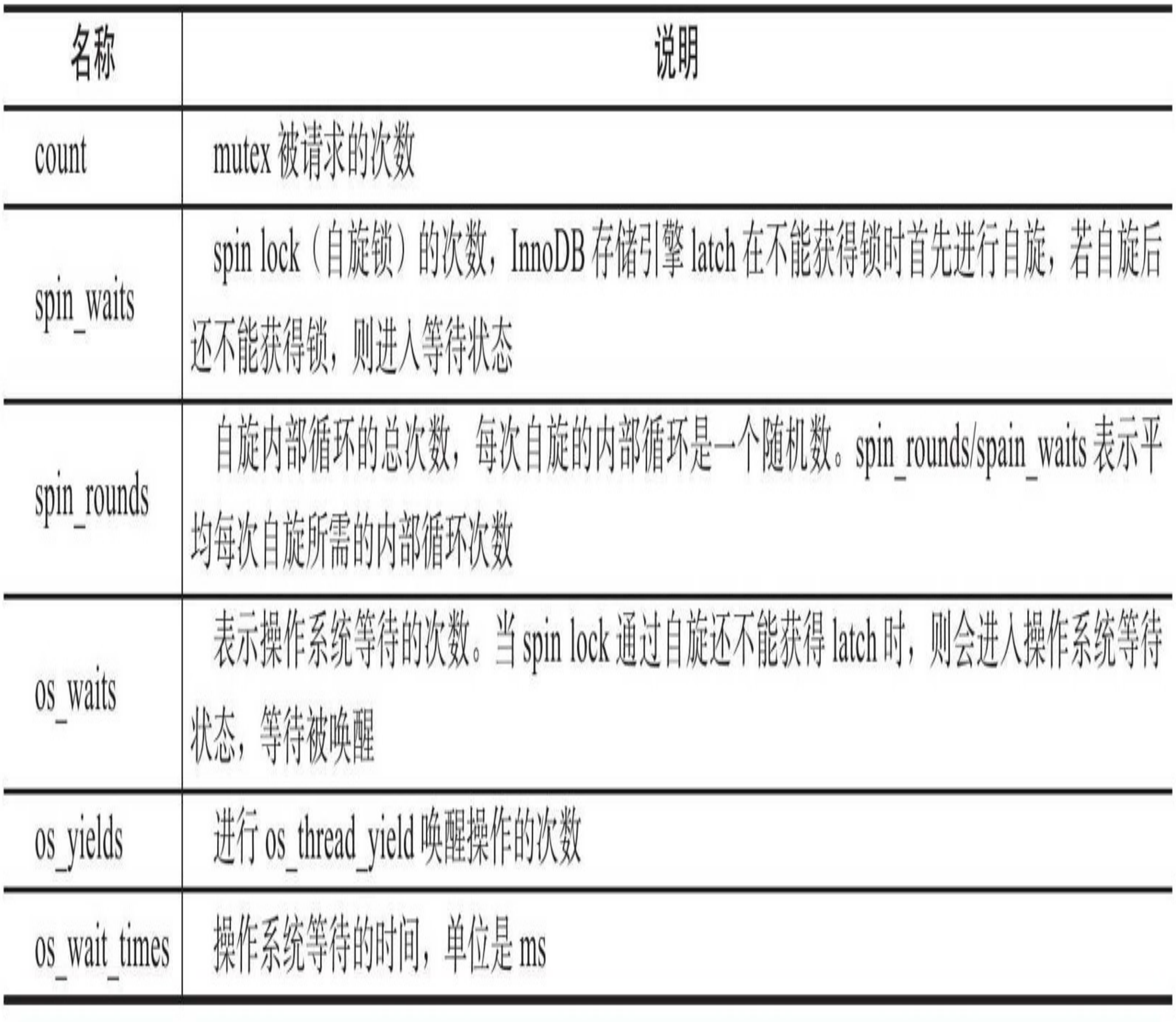


```
mysql> SHOW ENGINE INNODB MUTEX;
+--------+----------------------+---------------------------------------------------------------------------------+
| Type   | Name                 | Status                                                                          |
+--------+----------------------+---------------------------------------------------------------------------------+
| InnoDB | &kernel_mutex:srv0srv.c | count=54, spin_waits=6, spin_rounds=60, os_waits=3, os_yields=3, os_wait_times=0 |
| InnoDB | log0log.c:833        | os_waits=2                                                                      |
| InnoDB | rw_lock_mutexes      | count=0, spin_waits=0, spin_rounds=0, os_waits=0, os_yields=0, os_wait_times=0  |
+--------+----------------------+---------------------------------------------------------------------------------+
3 rows in set (0.01 sec)
```

□ 6-2 □Debug□□□□□□□latch

□□□□□□□□□□□□□□Type□□□□□InnoDB□□Name□□□□latch□□□□□□□□□□□□□□□□□□Status□□□□□□Debug□□□□□□□□os_waits□□□□□count□spin_waits□spin_rounds□os_yields□os_wait_times□□□□□□□□□□□6-2□

表 6-2 命令 SHOW ENGINE INNODB MUTEX 输出结果说明



名称	说明
count	mutex 被请求的次数
spin_waits	spin lock（自旋锁）的次数，InnoDB 存储引擎 latch 在不能获得锁时首先进行自旋，若自旋后还不能获得锁，则进入等待状态
spin_rounds	自旋内部循环的总次数，每次自旋的内部循环是一个随机数。spin_rounds/spain_waits 表示平均每次自旋所需的内部循环次数
os_waits	表示操作系统等待的次数。当 spin lock 通过自旋还不能获得 latch 时，则会进入操作系统等待状态，等待被唤醒
os_yields	进行 os_thread_yield 唤醒操作的次数
os_wait_times	操作系统等待的时间，单位是 ms

□□

[illegible]latch[illegible]lock[illegible]SHOW ENGINE INNODB STATUS[illegible]information_schema[illegible]INNODB_TRX[illegible]INNODB_LOCKS[illegible]INNODB_LOCK_WAITS[illegible]

6.3 InnoDB□□□□□□□

6.3.1 □□□□

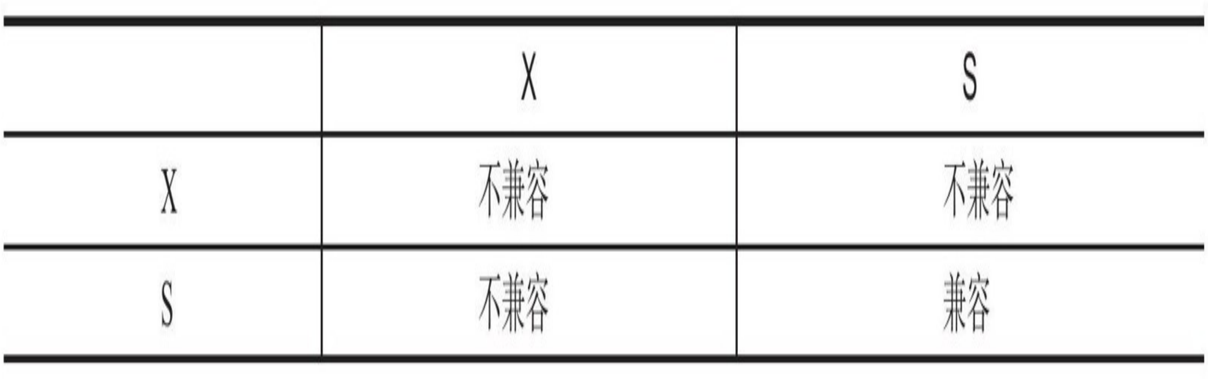
InnoDB□□□□□□□□□□□□□□□□□□

❏□□□□S Lock□□□□□□□□□□□□

❏□□□□X Lock□□□□□□□□□□□□□□□□

□□□□□□T1□□□□□□r□□□□□□□□□□□□T2□□□□□□□r□□□□□□□□□□□□□□□r□□□□□□□□□□□□□□Lock Compatible□□□□□□□□□□T3□□□□r□□□□□□□□□□□□□T1□T2□□□r□□□□□——□□□□□□□□□□□□6-3□□□□□□□□□□□□□□□

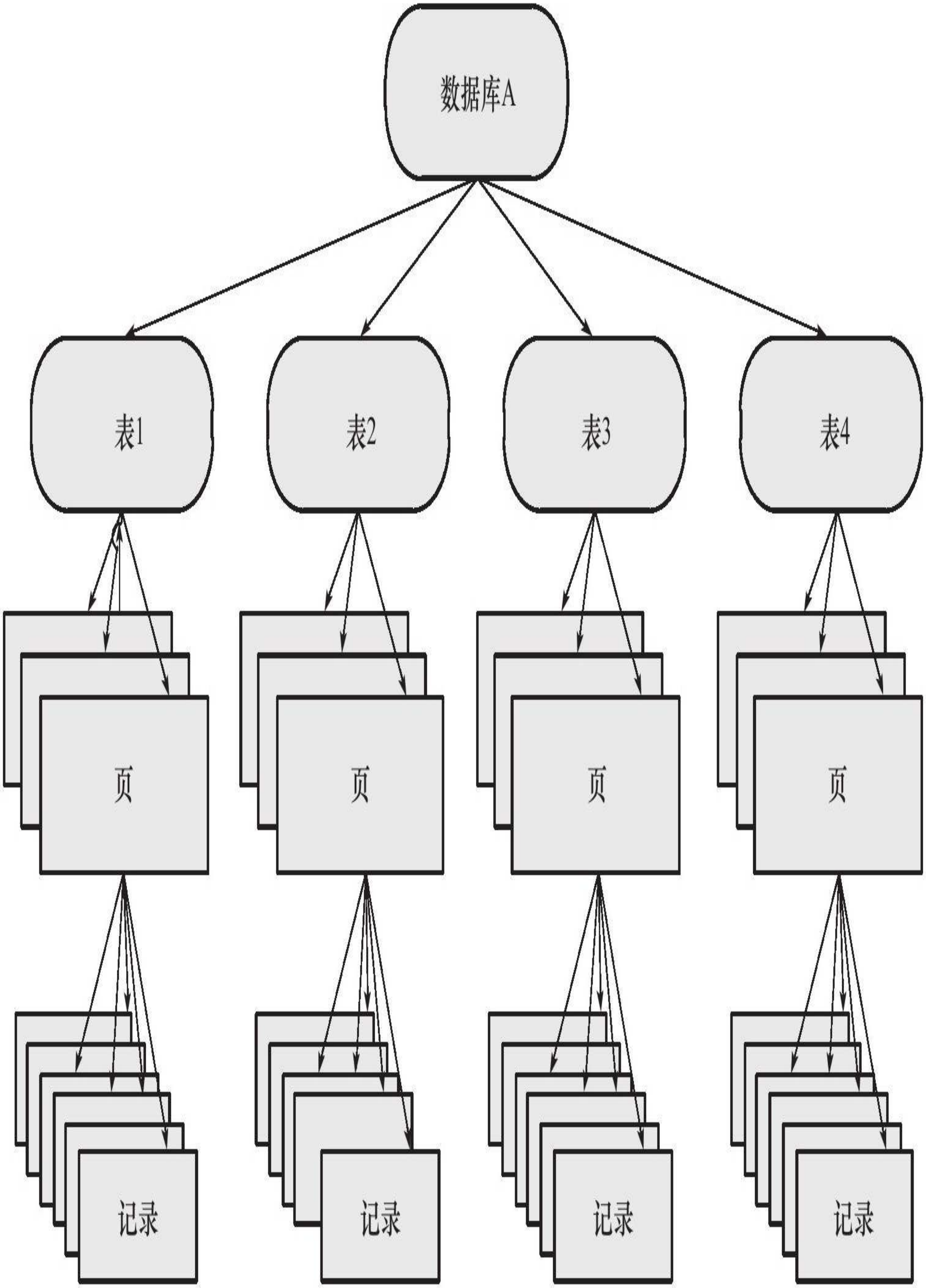
表6-3 排他锁和共享锁的兼容性



	X	S
X	不兼容	不兼容
S	不兼容	兼容

□□6-3□□□□X□□□□□□□□□□□□S□□□S□□□□□□□□□□□□□S□X□□□□□□□□□□□□□□□□row□□□□□□□□□

□□□InnoDB□□□□□□□□□□granular□□InnoDB□□□□□□□□□□□□□□□□□□□□□□Intention Lock□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□fine granularity□□□□□□□□□6-3□□□


数据库A
表1
表2
表3
表4
页
页
页
页
记录
记录
记录
记录

□ 6-3 □□□□

□□□6-3□□□□□□□□□□□r□□□X□□□□□□□□□□□□A□□□□□□□□IX□□□□□□r□X□□□r□X□□□□□□□□□□□1□□□S□□□□□□1□□□□S□□□□□□□□□□□r□□1□□□IX□□□□□□□□□□□□□□□□□□□□□□□□

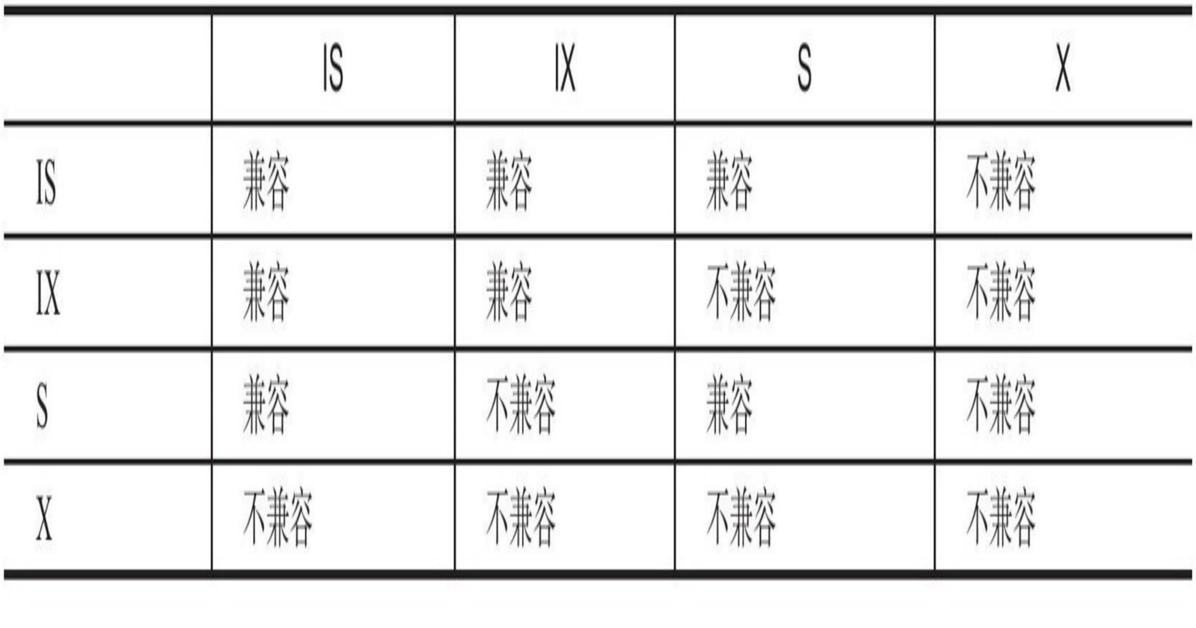
InnoDB□□□

1□□□□□□□IS Lock□□□□□□□□□□□□□□□□□□□

2□□□□□□□IX Lock□□□□□□□□□□□□□□□□□□□

□□InnoDB□□6-4□□□

表 6-4 InnoDB 存储引擎中锁的兼容性



	IS	IX	S	X
IS	兼容	兼容	兼容	不兼容
IX	兼容	兼容	不兼容	不兼容
S	兼容	不兼容	兼容	不兼容
X	不兼容	不兼容	不兼容	不兼容

□□□□□□□□SHOW ENGINE INNODB STATUS□□□□□□□□□□□□□□

```
mysql＞SHOW ENGINE INNODB STATUS\G;

……

------------

TRANSACTIONS

------------

Trx id counter 48B89BF

Purge done for trx's n:o＜48B89BA undo n:o＜0

History list length 0

LIST OF TRANSACTIONS FOR EACH SESSION:

---TRANSACTION 0,not started,process no 13757,OS thread id 1255176512

MySQL thread id 42,query id 80424887 localhost root

show engine innodb status

---TRANSACTION 48B89BE,ACTIVE 193 sec,process no 13757,OS thread id 1254910272 starting index read

mysql tables in use 1,locked 1

LOCK WAIT 2 lock struct(s),heap size 368,1 row lock(s)

MySQL thread id 41,query id 80424886 localhost root Sending data

select*from t where a＜4 lock in share mode

-------TRX HAS BEEN WAITING 2 SEC FOR THIS LOCK TO BE GRANTED:

RECORD LOCKS space id 30 page no 3 n bits 72 index'PRIMARY'of table'test'.'t'trx id 48B89BE lock mode S waiting

------------------

TABLE LOCK table'test'.'t'trx id 48B89BE lock mode IS

RECORD LOCKS space id 30 page no 3 n bits 72 index'PRIMARY'of table'test'.'t'trx id 48B89BE lock mode S waiting

---TRANSACTION 48B89BD,ACTIVE 205 sec,process no 13757,OS thread id 1257838912

2 lock struct(s),heap size 368,1 row lock(s)

MySQL thread id 40,query id 80424881 localhost root

TABLE LOCK table'test'.'t'trx id 48B89BD lock mode IX

RECORD LOCKS space id 30 page no 3 n bits 72 index'PRIMARY'of table'test'.'t'trx id 48B89BD lock_mode X locks rec but not gap

----------------------------

END OF INNODB MONITOR OUTPUT
```

```
===========================

1 row in set(0.01 sec)
```

□□□□SQL□□select*from t where a□4 lock in share mode□□□□RECORD LOCKS space id 30 page no 3 n bits 72 index'PRIMARY'of table'test'.'t'trx id 48B89BD lock_mode X locks rec but not gap□□□□□□□□locks rec but not gap□□□□□□□□□□□□□□□□□□

□InnoDB 1.0□□□□□□□□□□□□□SHOW FULL PROCESSLIST□SHOW ENGINE INNODB STATUS□□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB1.0□□□□INFORMATION_SCHEMA□□□□□□□INNODB_TRX□INNODB_LOCKS□INNODB_LOCK_WAITS□□□6-5□□INNODB_TRX□□□□□□8□□□□□□

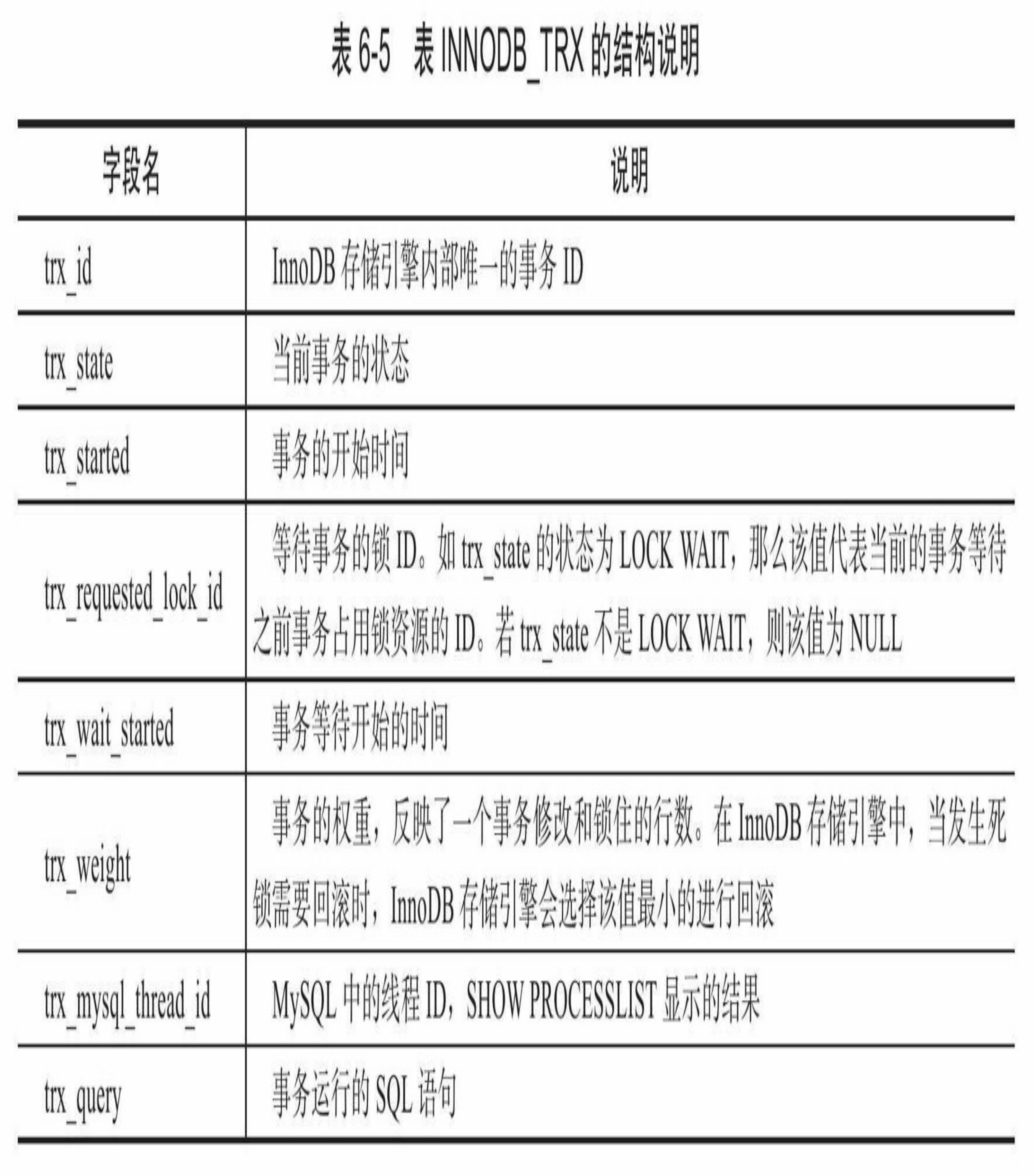


表6-5 表INNODB_TRX的结构说明

字段名	说明
trx_id	InnoDB存储引擎内部唯一的事务ID
trx_state	当前事务的状态
trx_started	事务的开始时间
trx_requested_lock_id	等待事务的锁ID。如trx_state的状态为LOCK WAIT，那么该值代表当前的事务等待之前事务占用锁资源的ID。若trx_state不是LOCK WAIT，则该值为NULL
trx_wait_started	事务等待开始的时间
trx_weight	事务的权重，反映了一个事务修改和锁住的行数。在InnoDB存储引擎中，当发生死锁需要回滚时，InnoDB存储引擎会选择该值最小的进行回滚
trx_mysql_thread_id	MySQL中的线程ID，SHOW PROCESSLIST显示的结果
trx_query	事务运行的SQL语句

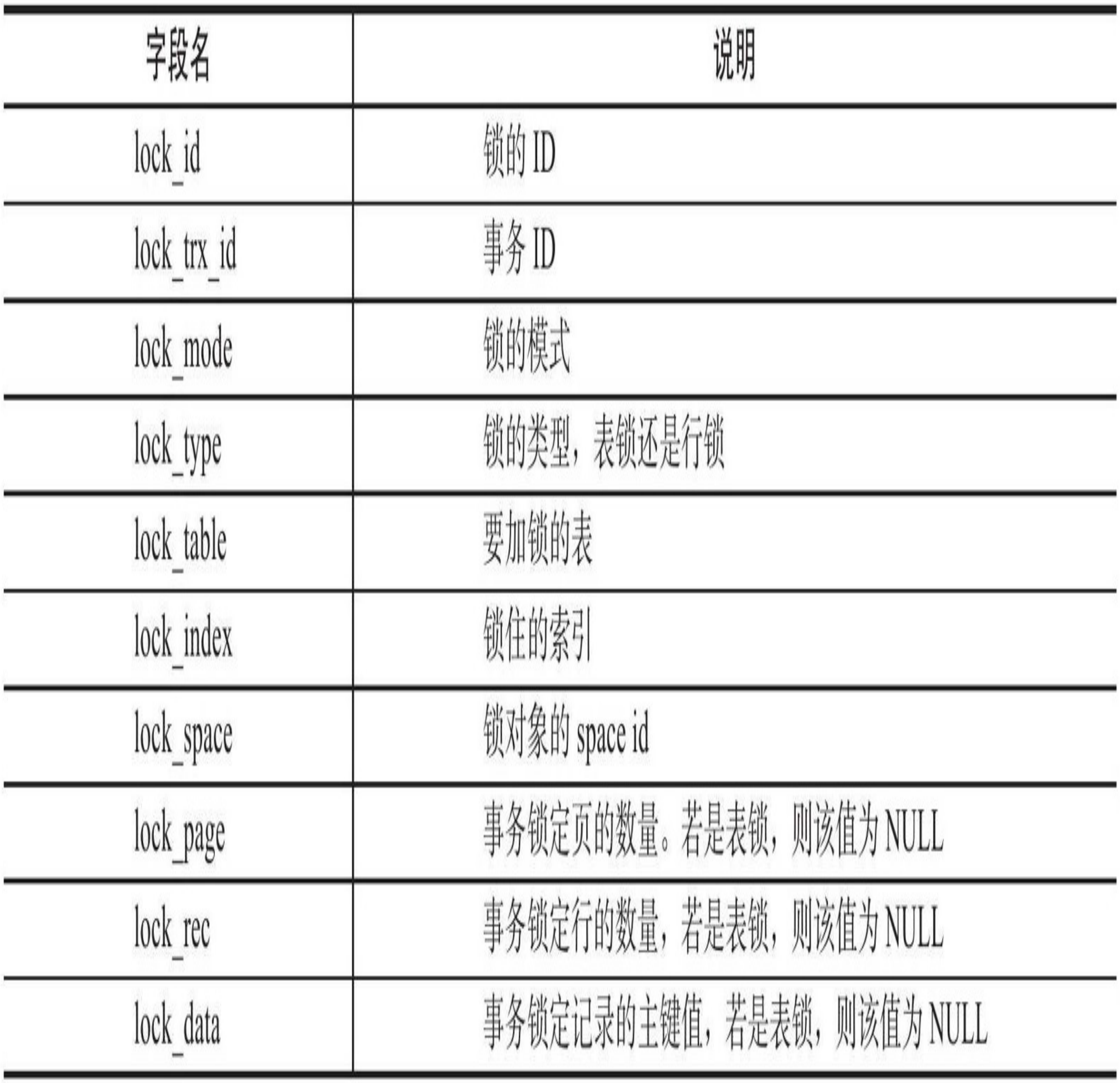
□□□□□□□□□□□□

```
mysql□SELECT*FROM information_schema.INNODB_TRX\G;

***************************1.row***************************

trx_id:7311F4
```

```
trx_state:LOCK WAIT
trx_started:2010-01-04 10:49:33
trx_requested_lock_id:7311F4:96:3:2
trx_wait_started:2010-01-04 10:49:33
trx_weight:2
trx_mysql_thread_id:471719
trx_query:select*from parent lock in share mode
***************************2.row***************************
trx_id:730FEE
trx_state:RUNNING
trx_started:2010-01-04 10:18:37
trx_requested_lock_id:NULL
trx_wait_started:NULL
trx_weight:2
trx_mysql_thread_id:471718
trx_query:NULL
2 rows in set(0.00 sec)
```

□□□state□□□□□trx_id□730FEE□□□□□□□□□□□trx_id□7311F4□□□□□□□“LOCK WAIT”□□□□□□□SQL□□□select*from parent lock in share mode□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□INNODB_LOCKS□□□□□□□□□□6-6□□□

表6-6 表INNODB_LOCKS的结构



字段名	说明
lock_id	锁的ID
lock_trx_id	事务ID
lock_mode	锁的模式
lock_type	锁的类型，表锁还是行锁
lock_table	要加锁的表
lock_index	锁住的索引
lock_space	锁对象的space id
lock_page	事务锁定页的数量。若是表锁，则该值为NULL
lock_rec	事务锁定行的数量，若是表锁，则该值为NULL
lock_data	事务锁定记录的主键值，若是表锁，则该值为NULL

□□□□□□□□□□□□□INNODB_LOCKS□

```
mysql□SELECT*FROM information_schema.INNODB_LOCKS\G;

***************************1.row***************************

lock_id:7311F4:96:3:2

lock_trx_id:7311F4
```

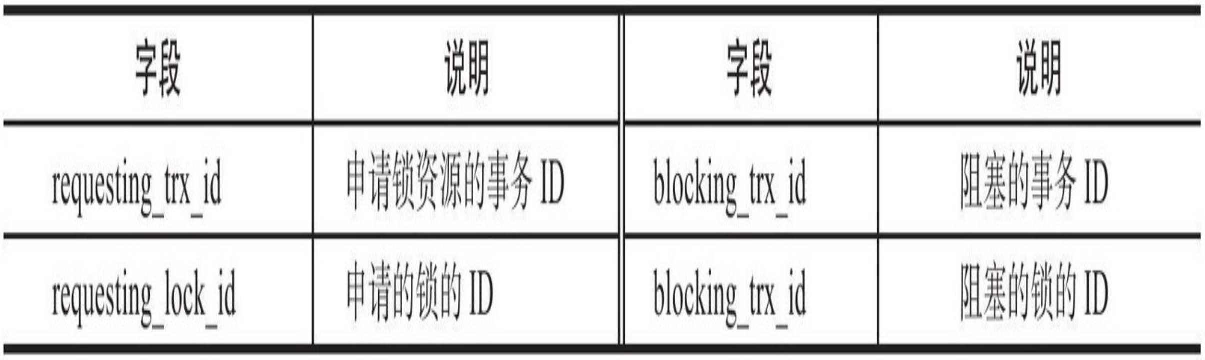
```
lock_mode:S

lock_type:RECORD

lock_table:'mytest'.'parent'

lock_index:'PRIMARY'

lock_space:96

lock_page:3

lock_rec:2

lock_data:1

***************************2.row***************************

lock_id:730FEE:96:3:2

lock_trx_id:730FEE

lock_mode:X

lock_type:RECORD

lock_table:'mytest'.'parent'

lock_index:'PRIMARY'

lock_space:96

lock_page:3

lock_rec:2

lock_data:1

2 rows in set(0.00 sec)
```

□□□□□□□□□□□□□□□□□□trx_id□730FEE□□□□□parent□□□□X□□□□ID□7311F4□□□□□parent□□□□□S□□□□lock_data□□1□□□□□□□□□□□□□□□□□□□□□□INNODB_TRX□□□□□□□□□trx_state□“RUNNING”□□□□□“LOCK WAIT”□□

□□□□□□□□□□□□□□lock_data□□□□□□“□□”□□□□□□□□□□□□□□□□□□lock_data□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□INNODB_LOCKS□□□□□□□□□□□NULL□□InnoDB□□□□□□□□□□□□□□□□□□

[illegible]INNODB_LOCKS[illegible]INNODB_LOCK_WAITS[illegible]INNODB_LOCK_WAITS[illegible]4[illegible]6-7[illegible]

表 6-7 表 INNODB_LOCK_WAITS 的结构



字段	说明	字段	说明
requesting_trx_id	申请锁资源的事务 ID	blocking_trx_id	阻塞的事务 ID
requesting_lock_id	申请的锁的 ID	blocking_trx_id	阻塞的锁的 ID

[illegible]

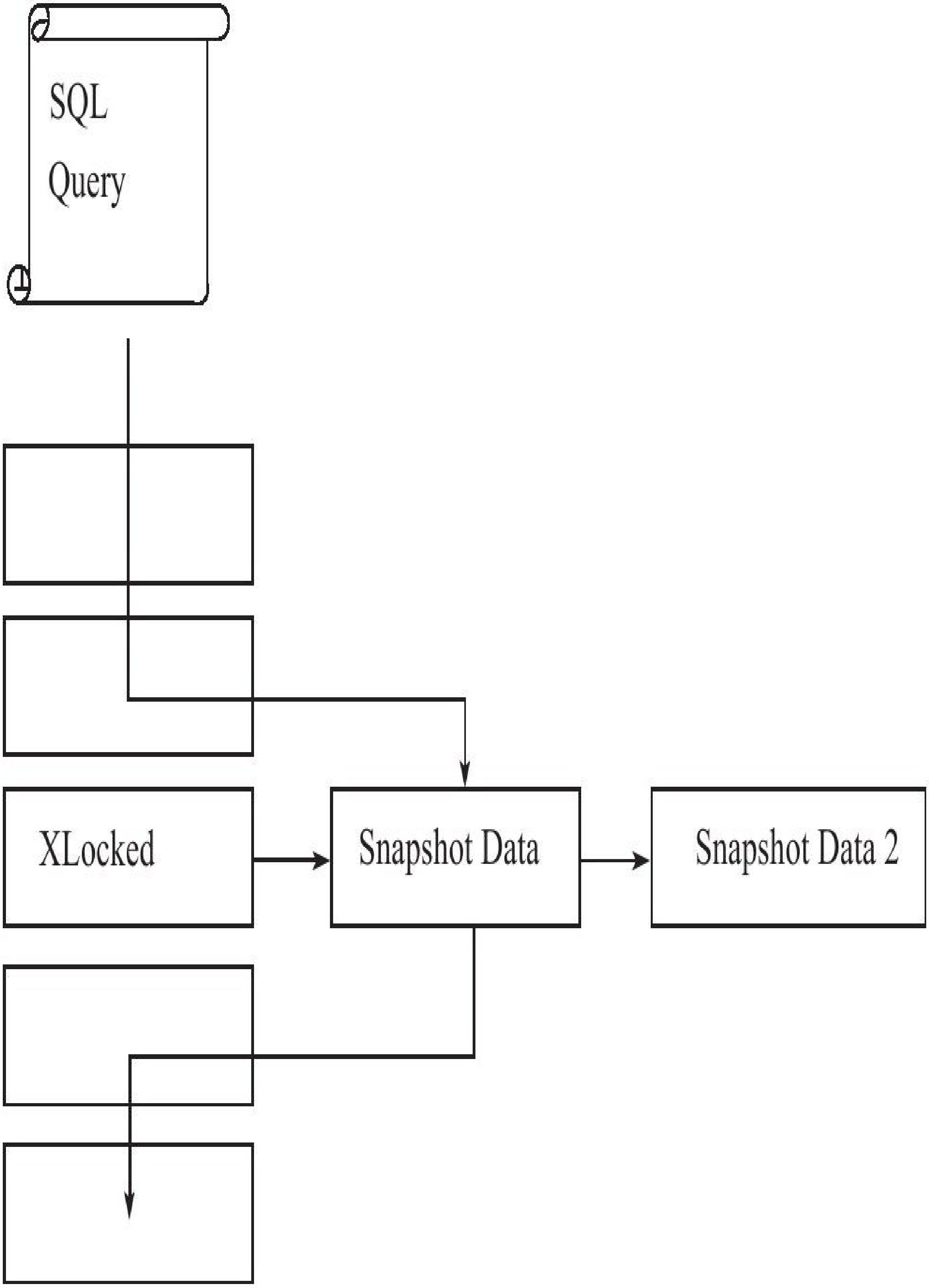
```
mysql□SELECT*FROM information_schema.INNODB_LOCK_WAITS\G;

***************************1.row***************************

requesting_trx_id:7311F4

requested_lock_id:7311F4:96:3:2

blocking_trx_id:730FEE

blocking_lock_id:730FEE:96:3:2

1 row in set(0.00 sec)
```

[illegible]SQL[illegible]ID[illegible]INNODB_TRX[illegible]INNODB_LOCKS[illegible]INNODB_LOCK_WAITS[illegible]

```
mysql□SELECT

r.trx_id waiting_trx_id,
r.trx_mysql_thread_id waiting_thread,
r.trx_query waiting_query,
b.trx_id blocking_trx_id,
b.trx_mysql_thread_id blocking_thread,
b.trx_query blocking_query
FROM information_schema.innodb_lock_waits w
INNER JOIN information_schema.innodb_trx b
ON b.trx_id=w.blocking_trx_id
INNER JOIN information_schema.innodb_trx r
ON r.trx_id=w.requesting_trx_id\G;
***************************1.row***************************
waiting_trx_id:73122F
waiting_thread:471719
waiting_query:NULL
blocking_trx_id:7311FC
blocking_thread:471718
blocking_query:NULL
1 row in set(0.00 sec)
```

6.3.2 □□□□□□□

□□□□□□□□□consistent nonlocking read□□□InnoDB□□□□□□□□□□□□□multi versioning□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□DELETE□UPDATE□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□6-4□□□


SQL
Query
XLocked
Snapshot Data
Snapshot Data 2

□ 6-4 InnoDB□□□□□□□□□□□□

□6-4□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□X□□□□□□□□□□□□□□□□□□□□□□□□□□□□undo□□□□□□undo□□

□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□

□□□6-4□□6-4□□Multi Version Concurrency Control□MVCC□□

□□□□□□□READ COMMITTED□REPEATABLE READ□InnoDB□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□READ COMMITTED□□REPEATABLE READ□□MySQL□□□□□□□□A□□□□□SQL□□□

```
#Session A

mysql□BEGIN;

Query OK,0 rows affected(0.00 sec)

mysql□SELECT*FROM parent WHERE id=1;

+----+

|id|

+----+

|1|

+----+

1 row in set(0.00 sec)
```

□□A□□□□□□□□□□□BEGIN□□□□□□□□□□□□□parent□id□1□□□□□□□□□□□□□□□□□□□□□□□□□□□□□B□□□□□□□□□□□□□□□□□□B□□□□□□□

```
mysql□BEGIN;

Query OK,0 rows affected(0.00 sec)

mysql□UPDATE parent SET id=3 WHERE id=1;

Query OK,1 row affected(0.00 sec)

Rows matched:1 Changed:1 Warnings:0
```

□□□B□□□□□parent□id□1□□□□□□id=3□□□□□□□□□□□□□□id=1□□□□□□□□□X□□□□□□□□□A□□□□□id□1□□□□□□InnoDB□□□□□□□□□□READ COMMITTED□REPEATETABLE READ□□□□□□□□□□□□□□□□□□□□□□□□□□□A□□□□□□□□□□□□□□SQL□□SELECT*FROM parent WHERE id=1□□□□□□□□□□READ COMMITTED□□REPEATABLE READ□□□□□□□□□□□□□□□□□□

```
mysql□SELECT*FROM parent WHERE id=1;

+----+

|id|

+----+

|1|

+----+

1 row in set(0.00 sec)
```

□□□□id=1□□□□□□□1□□□□□□□□□□□□□□□□□□□□□B□□□□□□□□□

```
#Session B

mysql□commit;

Query OK,0 rows affected(0.01 sec)
```

□□□B□□□□□□□□□□□A□□□□SELECT*FROM parent WHERE id=1□SQL□□□□READ COMMITTED□REPEATABLE□□□□□□□□□□□□□□□□□□□READ COMMITTED□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□fresh snapshot□□□□□□□□□□□□□B□□□□□□□□□□READ COMMITTED□□□□□□□□□□□□□□□

```
mysql□SELECT@@tx_isolation\G;

***************************1.row***************************

@@tx_isolation:READ-COMMITTED

1 row in set(0.00 sec)

mysql□SELECT*FROM parent WHERE id=1;

Empty set(0.00 sec)
```

□□□REPEATABLE READ□□□□□□□□□□□□□□□□□□□□□□□□□□REPEATABLE READ□□□□□□□□□□□□□□□□

```
mysql□SELECT@@tx_isolation\G;

***************************1.row***************************

@@tx_isolation:REPEATABLE-READ

1 row in set(0.00 sec)

mysql□SELECT*FROM parent WHERE id=1;

+----+

|id|

+----+

|1|

+----+

1 row in set(0.00 sec)
```

□□□□□□□□□□□□□□□□□□□□□□□6-8□□□□□□□□□□□□□□READ COMMITTED□□□□□□□□□□□□□□□□□□□□□□□□□□□□ACID□□I□□□□□□□□□□□□□7□□□□□□□□□

表 6-8　示例执行的过程

时间	会话 A	会话 B
1	BEGIN	
2	SELECT * FROM parent WHERE id = 1;	
3		BEGIN
4		UPDATE parent SET id=3 WHERE id = 1;
5	SELECT * FROM parent WHERE id = 1;	
6		COMMIT;
7	SELECT * FROM parent WHERE id = 1;	
8	COMMIT	

6.3.3 [illegible]

[illegible]REPEATABLE READ[illegible]InnoDB [illegible]SELECT[illegible]SELECT[illegible]InnoDB [illegible]SELECT[illegible]locking read[illegible]

❑SELECT…FOR UPDATE

❑SELECT…LOCK IN SHARE MODE

SELECT…FOR UPDATE[illegible]X[illegible]SELECT…LOCK IN SHARE MODE[illegible]S[illegible]S[illegible]X[illegible]

[illegible]SELECT…FOR UPDATE[illegible]SELECT…FOR UPDATE[illegible]SELECT…LOCK IN SHARE MODE[illegible]SELECT[illegible]BEGIN[illegible]START TRANSACTION[illegible]SET AUTOCOMMIT=0[illegible]

6.3.4 [illegible]

[illegible]DBA[illegible]InnoDB[illegible]auto-increment counter[illegible]

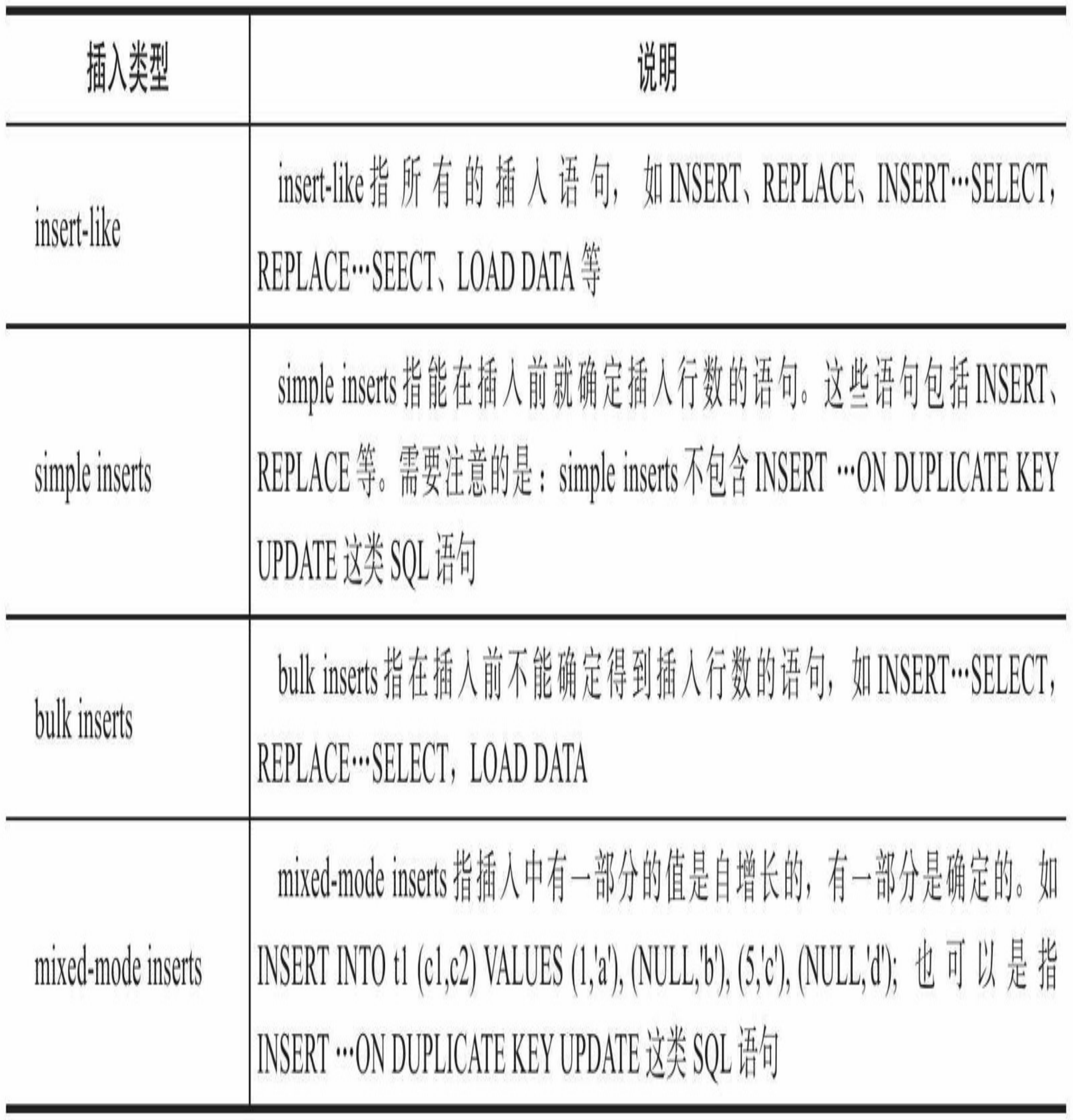
```
SELECT MAX(auto_inc_col)FROM t FOR UPDATE;
```

[illegible]1[illegible]AUTO-INC Locking[illegible]SQL[illegible]

[illegible]AUTO-INC Locking[illegible]INSERT…SELECT[illegible]

[illegible]MySQL 5.1.22[illegible]InnoDB[illegible]InnoDB[illegible]innodb_autoinc_lock_mode[illegible]1[illegible]6-9[illegible]

表 6-9 插入类型



插入类型	说明
insert-like	insert-like 指所有的插入语句，如 INSERT、REPLACE、INSERT…SELECT，REPLACE…SEECT、LOAD DATA 等
simple inserts	simple inserts 指能在插入前就确定插入行数的语句。这些语句包括 INSERT、REPLACE 等。需要注意的是：simple inserts 不包含 INSERT …ON DUPLICATE KEY UPDATE 这类 SQL 语句
bulk inserts	bulk inserts 指在插入前不能确定得到插入行数的语句，如 INSERT…SELECT，REPLACE…SELECT，LOAD DATA
mixed-mode inserts	mixed-mode inserts 指插入中有一部分的值是自增长的，有一部分是确定的。如 INSERT INTO t1 (c1,c2) VALUES (1,'a'), (NULL,'b'), (5,'c'), (NULL,'d'); 也可以是指 INSERT …ON DUPLICATE KEY UPDATE 这类 SQL 语句

□□□□□□□innodb_autoinc_lock_mode□□□□□□□□□□□□□□□□□□□□□□□□□□□□0□1□2□□□□□□□6-10□□□

表 6-10 参数 innodb_autoinc_lock_mode 的说明

nnodb_autoinc_lock_mode	说明
0	这是 MySQL5.1.22 版本之前自增长的实现方式，即通过表锁的 AUTO-INC Locking 方式。因为有了新的自增长实现方式，0 这个选项不应该是新版用户的首选项
1	这是该参数的默认值。对于“simple inserts”，该值会用互斥量（mutex）去对内存中的计数器进行累加的操作。对于“bulk inserts”，还是使用传统表锁的 AUTO-INC Locking 方式。在这种配置下，如果不考虑回滚操作，对于自增值列的增长还是连续的。并且在这种方式下，statement-based 方式的 replication 还是能很好地工作。需要注意的是，如果已经使用 AUTO-INC Locing 方式去产生自增长的值，而这时需要再进行“simple inserts”的操作时，还是需要等待 AUTO-INC Locking 的释放
2	在这个模式下，对于所有“INSERT-like”自增长值的产生都是通过互斥量，而不是 AUTO-INC Locking 的方式。显然，这是性能最高的方式。然而，这会带来一定的问题。因为并发插入的存在，在每次插入时，自增长的值可能不是连续的。此外，最重要的是，基于 Statement-Base Replication 会出现问题。因此，使用这个模式，任何时候都应该使用 row-base replication。这样才能保证最大的并发性能及 replication 主从数据的一致

□□□□□□□□□□□□InnoDB□□□□□□□□□□□□MyISAM□□□MyISAM□□□□□□□□□□□□□□□□□□□□□□□□□□□□master□□InnoDB□□□□□□slave□□MyISAM□□□□□replication□□□□□□□□□□□□□□□

□□□□InnoDB□□□MySQL□□□□□□□□□□MyISAM□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

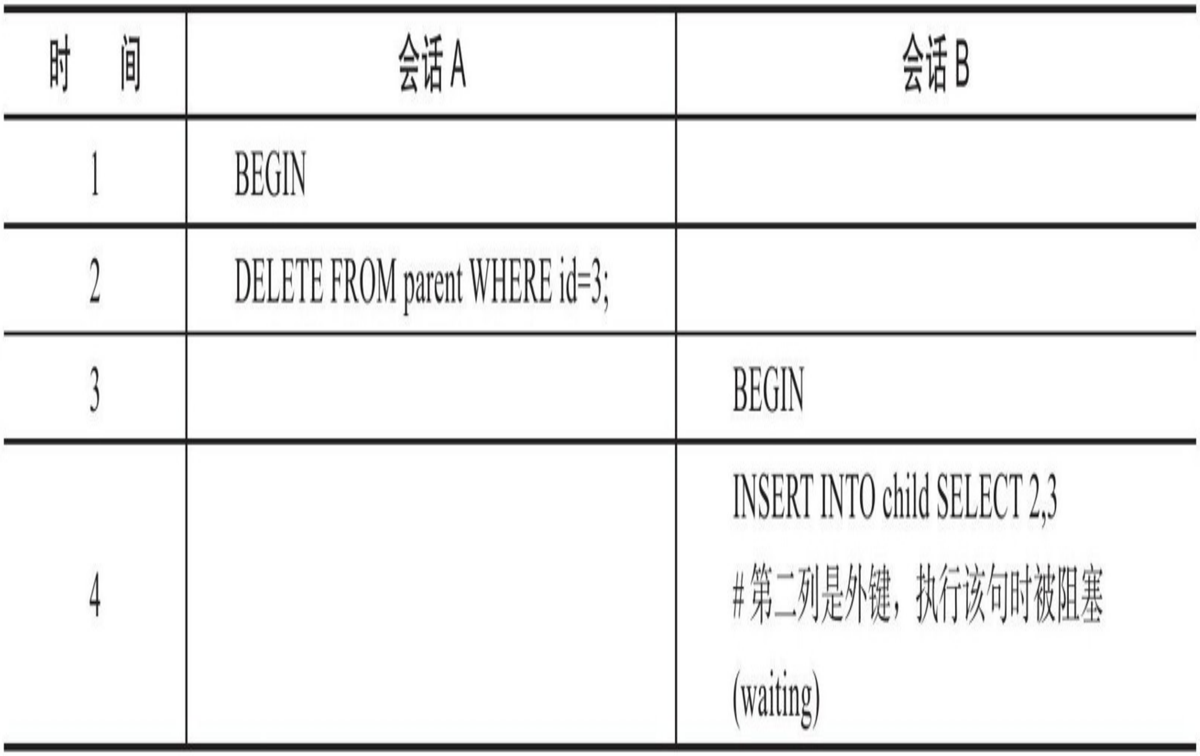
```
mysql□CREATE TABLE t(

-□a INT AUTO_INCREMENT,

-□B INT,

-□KEY(b,a)

-□)ENGINE=InnoDB;

ERROR 1075(42000):Incorrect table definition;there can be only one auto column and it must be defined as a key

mysql□CREATE TABLE t(

-□a INT AUTO_INCREMENT,

-□B INT,

-□KEY(b,a)

-□)ENGINE=MyISAM;

Query OK,0 rows affected(0.01 sec)
```

6.3.5 □□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□——□□Oracle□□□□□□□Oracle□□□□□□□□□□□□□□□□□□□□□□□□□□□□Oracle□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□SELECT□□□□□□□□□□SELECT□□□SELECT… LOCK IN SHARE MODE□□□□□□□□□□□□S□□□□□□□□□□□□□□X□□□□□□□□□□□□□□□6-11□□□

表6-11 外键测试用例



时　间	会话A	会话B
1	BEGIN	
2	DELETE FROM parent WHERE id=3;	
3		BEGIN
4		INSERT INTO child SELECT 2,3 #第二列是外键，执行该句时被阻塞 (waiting)

□□□□□□□□□□□□□□□□□□□□□COMMIT□ROLLBACK□□□□□□B□□□□□□□□□□□□id□3□□□□□□A□□□□□□□X□□□□□□□□B□□□□□□□□□□id□3□□□

[illegible]S[illegible]INSERT[illegible]Session B[illegible]id=3[illegible]A[illegible]id[illegible]3[illegible]INNODB_LOCKS[illegible]

```
mysql[illegible]SELECT*FROM information_schema.INNODB_LOCKS\G;

***************************1.row***************************

lock_id:7573B8:96:3:4

lock_trx_id:7573B8

lock_mode:S

lock_type:RECORD

lock_table:'mytest'.'parent'

lock_index:'PRIMARY'

lock_space:96

lock_page:3

lock_rec:4

lock_data:3

***************************2.row***************************

lock_id:7573B3:96:3:4

lock_trx_id:7573B3

lock_mode:X

lock_type:RECORD

lock_table:'mytest'.'parent'

lock_index:'PRIMARY'

lock_space:96

lock_page:3

lock_rec:4

lock_data:3

2 rows in set(0.00 sec)
```

6.4 □□□□

6.4.1 □□□3□□□

InnoDB□□□□□3□□□□□□□□□□□□

❑Record Lock□□□□□□□□□

❑Gap Lock□□□□□□□□□□□□□□□□□□□□

❑Next-Key Lock:Gap Lock+Record Lock□□□□□□□□□□□□□□□□

Record Lock□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□

Next-Key Lock□□□□Gap Lock□Record Lock□□□□□□□□□Next-Key Lock□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□10□11□13□20□□□□□□□□□□□□□Next-Key Locking□□□□□

```
(-∞,10]

(10,11]

(11,13]

(13□20]

(20,+∞)
```

□□Next-Key Lock□□□□□□□Next-Key Locking□□□□□□□□□□□□Phantom Problem□□□predict lock□□□□□□□□□next-key locking□□□previous-key locking□□□□□□□□□□10□11□13□20□□□□previous-key locking□□□□□□□□□□□□□

```
(-∞,10)

[10,11)
```

[11,13)

[13□20)

[20,+∞)

□□□T1□□□□next-key locking□□□□□□□□

(10,11]□(11□13]

□□□□□□□12□□□□□□□□□□□□

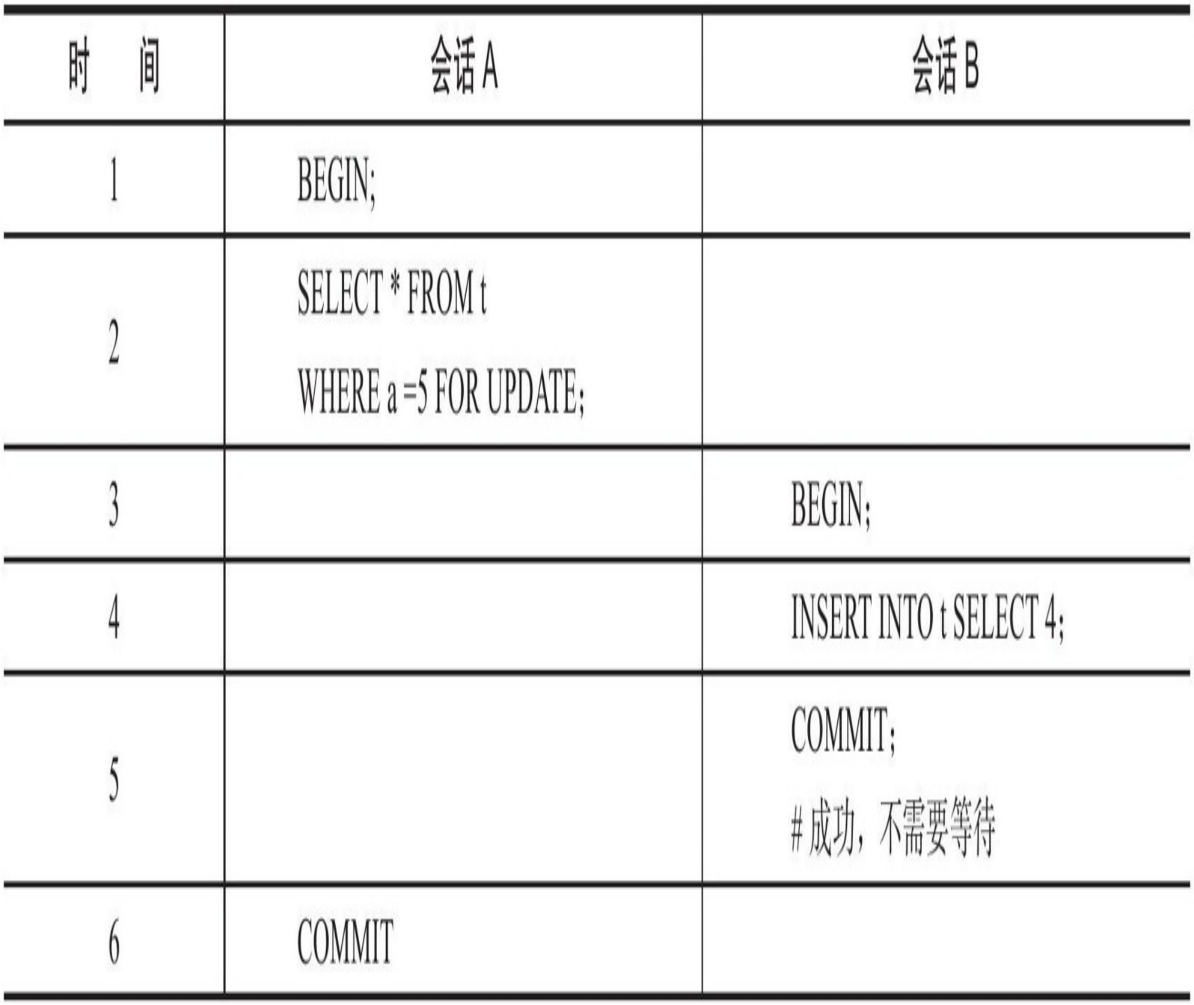
(10,11]□(11,12]□(12□13]

□□□□□□□□□□□□□□□□□InnoDB□□□□□□Next-Key Lock□□□□□□□□□□Record Lock□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□t□

```
DROP TABLE IF EXISTS t;

CREATE TABLE t(a INT PRIMARY KEY);

INSERT INTO t SELECT 1;

INSERT INTO t SELECT 2;

INSERT INTO t SELECT 5;
```

□□□□□□6-12□□SQL□□□

表6-12 唯一索引的锁定示例



时　间	会话A	会话B
1	BEGIN;	

2	SELECT * FROM t
	WHERE a =5 FOR UPDATE;	

3		BEGIN;
4		INSERT INTO t SELECT 4;
5		COMMIT; # 成功，不需要等待
6	COMMIT	

□t□□1□2□5□□□□□□□□□□□□□□□A□□□□a=5□□X□□□□□□a□□□□□□□□□□□□□□5□□□□□□□(2□5)□□□□□□□□□□B□□□□4□□□□□□□□□□□□□□□□□□□□Next-Key Lock□□□□□□Record Lock□□□□□□□□□□□□

□□□□□□□□□Next-Key Lock□□□Record Lock□□z□

```
CREATE TABLE z(a INT,b INT,PRIMARY KEY(a),KEY(b));

INSERT INTO z SELECT 1,1;
INSERT INTO z SELECT 3,1;

INSERT INTO z SELECT 5,3;

INSERT INTO z SELECT 7,6;

INSERT INTO z SELECT 10,8;
```

□z□□b□□□□□□□□□□A□□□□□□SQL□□□

```
SELECT*FROM z WHERE b=3 FOR UPDATE
```

□□□□□□SQL□□□□□□□b□□□□□□□□□□□□□Next-Key Locking□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□a□□5□□□□□Record Lock□□□□□□□□□□□□□□Next-Key Lock□□□□□□□(1□3)□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□gap lock□□□□□□□□□□□□□(3□6)□□□□□□□□□□□B□□□□□□SQL□□□□□□□□□

```
SELECT*FROM z WHERE a=5 LOCK IN SHARE MODE;

INSERT INTO z SELECT 4,2;

INSERT INTO z SELECT 6,5;
```

□□□SQL□□□□□□□□□□□□A□□□□SQL□□□□□□□□□□□a=5□□□□X□□□□□□□□□□□□□□SQL□□□□□□□4□□□□□□□□□□□□□□□□2□□□□□□(1□3)□□□□□□□□□□□□□□□□□SQL□□□□□□□□6□□□□□□5□□□□□(1□3)□□□□□□□□5□□□□□□□□□(3□6)□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□□□

```
INSERT INTO z SELECT 8,6;

INSERT INTO z SELECT 2,0;

INSERT INTO z SELECT 6,7;
```

□□□□□□□□□□□□Gap Lock□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Phantom Problem□□□□□□□□□□□□□□□□□□A□□□□□□□□b=3□□□□□□

[illegible]Gap Lock[illegible]3[illegible]6[illegible]b[illegible]3[illegible]A[illegible]Phantom Problem[illegible]

[illegible]Gap Lock[illegible]

❑[illegible]READ COMMITTED

❑[illegible]innodb_locks_unsafe_for_binlog[illegible]1

[illegible]Gap Lock[illegible]Record Lock[illegible]replication[illegible]READ COMMITTED[illegible]READ REPEATABLE[illegible]

[illegible]InnoDB[illegible]INSERT[illegible]A[illegible]z[illegible]b=3[illegible](1[illegible]3)[illegible]

```
INSERT INTO z SELECT 2,2;
```

[illegible]b[illegible]2[illegible]3[illegible]

```
INSERT INTO z SELECT 2,0;
```

[illegible]Next-Key Lock[illegible]Record Lock[illegible]range[illegible]point[illegible]InnoDB[illegible]Next-Key Lock[illegible]

6.4.2 [illegible]Phantom Problem

[illegible]REPEATABLE READ[illegible]InnoDB[illegible]Next-Key Locking[illegible]Phantom Problem[illegible]Oracle[illegible]SERIALIZABLE[illegible]Phantom Problem[illegible]

Phantom Problem[illegible]SQL[illegible]SQL[illegible]t[illegible]t[illegible]1[illegible]2[illegible]5[illegible]T1[illegible]SQL[illegible]

```
SELECT*FROM t WHERE a[illegible]2 FOR UPDATE;
```

[illegible]T1[illegible]5[illegible]T2[illegible]4[illegible]T1[illegible]SQL[illegible]4[illegible]5[illegible]6-13[illegible]

表 6-13　Phantom Problem 的演示

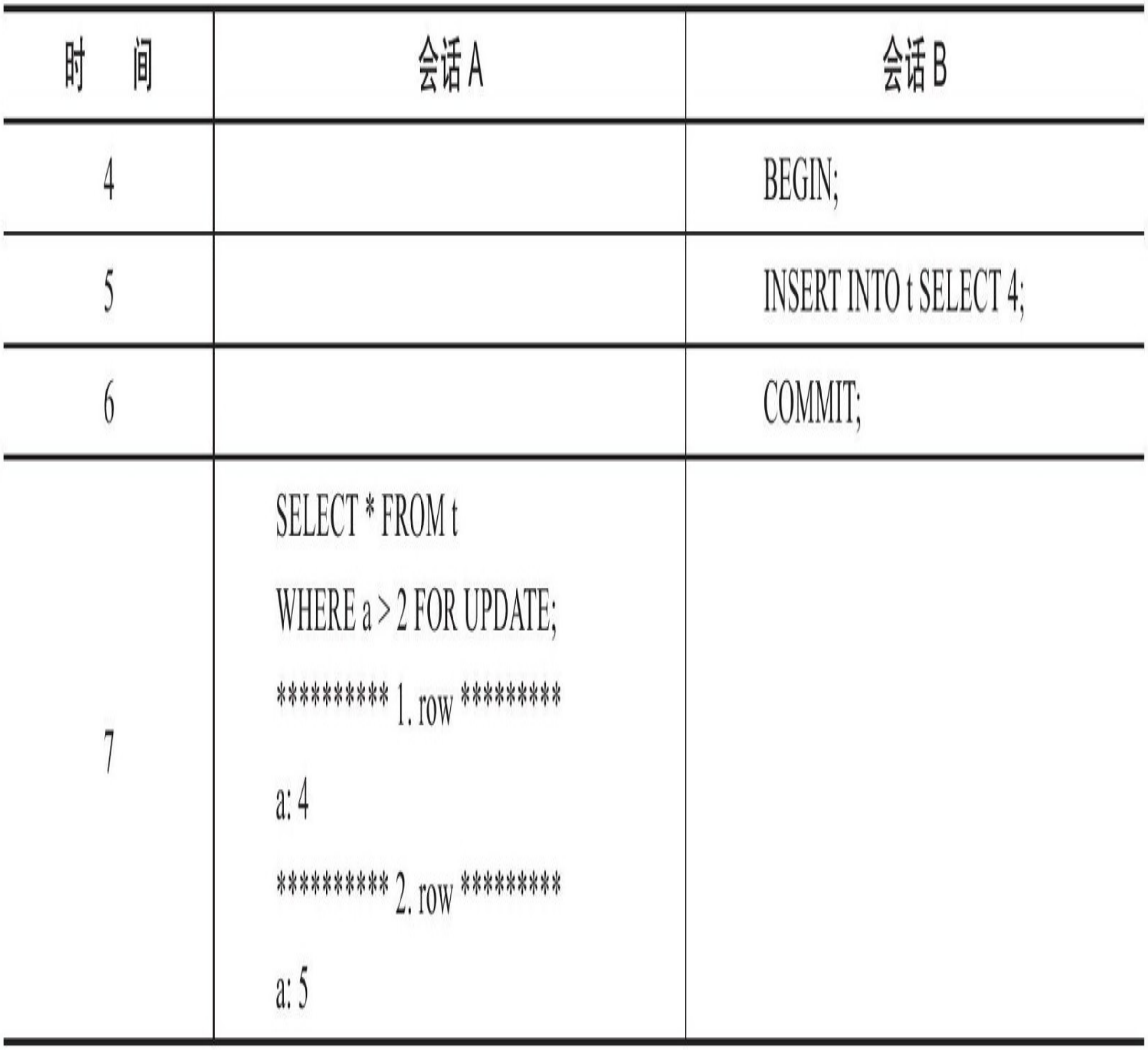
时　间	会话 A	会话 B

1	SET SESSION
	tx_isolation='READ-OMMITTED';	

2	BEGIN;	

3	SELECT * FROM t
	WHERE a > 2 FOR UPDATE;
	********* 1. row *********
	a: 4	

（续）



时　间	会话A	会话B
4		BEGIN;
5		INSERT INTO t SELECT 4;
6		COMMIT;
7	SELECT * FROM t WHERE a > 2 FOR UPDATE; ********* 1. row ********* a: 4 ********* 2. row ********* a: 5	

InnoDB[illegible]Next-Key Locking[illegible]Phantom Problem[illegible]SQL[illegible]SELECT*FROM t WHERE a[illegible]2 FOR UPDATE[illegible]5[illegible]2[illegible]+∞[illegible]X[illegible]Phantom Problem[illegible]

InnoDB[illegible]REPEATABLE READ[illegible]Next-Key Locking[illegible]READ COMMITTED[illegible]

Record Lock[illegible]A[illegible]READ COMMITTED[illegible]

[illegible]InnoDB[illegible]Next-Key Locking[illegible]

```
SELECT*FROM table WHERE col=xxx LOCK IN SHARE MODE[illegible]

If not found any row:

#unique for insert value

INSERT INTO table VALUES(...);
```

[illegible]SLock[illegible]SELECT…LOCK IN SHARE MODE[illegible]6-14[illegible]

表6-14　通过Next-Key Locking实现应用程序的唯一性检查

时　间	会话A	会话B
1	BEGIN	

2	mysql>SELECT * FROM z
	WHERE b=4
	LOCK IN SHARE MODE;	

3		mysql>SELECT * FROM z
		WHERE b=4
		LOCK IN SHARE MODE;

4	mysql>INSERT INTO z SELECT 4,4; #阻塞	
5		mysql>INSERT INTO z SELECT4,4; ERROR 1213 (40001):Deadlock found when trying to get lock;try restarting transaction #抛出死锁异常
6	#INSERT插入成功	

6.5 [illegible]

[illegible]

6.5.1 [illegible]

[illegible]Dirty Read[illegible]commit[illegible]

[illegible]

[illegible]

[illegible]6-15[illegible]

表 6-15　脏读的示例

Time	会话 A	会话 B

1	SET
	@@tx_isolation='read-ncommitted';	

2		SET
		@@tx_isolation='read-ncommitted';

3		BEGIN;
4		mysql> SELECT * FROM t\G; ********* 1. row ************ a: 1 1 row in set (0.00 sec)
5	INSERT INTO t SELECT 2;	
6		mysql> SELECT * FROM t\G; ********* 1. row ************ a: 1 ********* 2. row ************ a: 2 2 row in set (0.00 sec)

□t□□□□□□6.4.1□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□REPEATABLE READ□□□READ UNCOMMITTED□□□□□□A□□□□□□□□□□□□□□□□□□B□□□□SELECT□□□□□□□□□□□□□2□□□□□□□□A□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□READ UNCOMMITTED□□□□□□□□□□□□□□□□□□READ COMMITTED□InnoDB□□□□□□□□□□□□□□READ REPEATABLE□Microsoft SQL Server□□□□READ COMMITTED□Oracle□□□□□□□READ COMMITTED□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□READ UNCOMMITTED□□□replication□□□□slave□□□□□□□slave□□□□□□□□□□□□□□□□□

6.5.2 [illegible]

[illegible]DML[illegible]

[illegible]6-16[illegible]

表 6-16　不可重复读的示例

Time	会话 A	会话 B
1	SET@@tx_isolation='read-committed';	
2		SET @@tx_isolation='read-committed';
3	BEGIN	BEGIN
4	mysql>SELECT * FROM t ; ********* 1. row ************ a: 1 1 row in set (0.00 sec)	
5		INSERT INTO t SELECT 2;
6		COMMIT;
7	mysql>SELECT * FROM t ; ********* 1. row ************ a: 1 ********* 1. row ************ a: 2 2 row in set (0.00 sec)	

[illegible]A[illegible]1[illegible]B[illegible]2[illegible]A[illegible]1[illegible]B[illegible]A[illegible]1[illegible]2[illegible]A[illegible]B[illegible]READ COMMITTED[illegible]

[illegible]Oracle[illegible]Microsoft SQL Server[illegible]READ COMMITTED[illegible]

[illegible]InnoDB[illegible]Next-Key Lock[illegible]MySQL[illegible]Phantom Problem[illegible]Next-Key Lock[illegible]gap[illegible]InnoDB[illegible]READ REPEATABLE[illegible]Next-Key Lock[illegible]

6.5.3 [illegible]

[illegible]

1[illegible]T1[illegible]r[illegible]v1[illegible]T1[illegible]

2[illegible]T2[illegible]r[illegible]v2[illegible]T2[illegible]

3[illegible]T1[illegible]

4[illegible]T2[illegible]

[illegible]READ UNCOMMITTED[illegible]DML[illegible]2[illegible]T2[illegible]r[illegible]T1[illegible]

[illegible]

1[illegible]T1[illegible]User1[illegible]

2[illegible]T2[illegible]User2[illegible]

3[illegible]User1[illegible]

4[illegible]User2[illegible]

[illegible]User1[illegible]“[illegible]”[illegible]“[illegible]”[illegible]10 000[illegible]9000[illegible]1[illegible]9999[illegible]9000[illegible]9000[illegible]USB Key[illegible]USB Key[illegible]

[illegible]1[illegible]X[illegible]2[illegible]X[illegible]2[illegible]1[illegible]3[illegible]4[illegible]6-17[illegible]

表 6-17　丢失更新问题的处理方法

Time	会话 A	会话 B
1	BEGIN;	

2	SELECT cash into @cash
	FROM account
	WHERE user = pUser FOR UPDATE;	

3		SELECT cash into @cash
		FROM account
		WHERE user = pUser FOR UPDATE;
		# 等待

	……	……

m	UPDATE account
	SET cash=@cash-9000
	WHERE user=pUser	

m+1	COMMIT	

m+2		UPDATE account SET cash=@cash-1
		WHERE user=pUser;

m+3		COMMIT

[illegible]UPDATE[illegible]SELECT…FOR UPDATE[illegible]UPDATE[illegible]UPDATE[illegible] SELECT[illegible]UPDATE[illegible]SQL[illegible]

[illegible]SELECT[illegible]INSERT[illegible]UPDATE[illegible]DELETE[illegible]

6.6 [illegible]

[illegible]

[illegible]InnoDB[illegible]innodb_lock_wait_timeout[illegible]50[illegible]innodb_rollback_on_timeout[illegible]OFF[illegible]innodb_lock_wait_timeout[illegible]MySQL[illegible]

```
mysql[illegible]SET@@innodb_lock_wait_timeout=60;

Query OK,0 rows affected(0.00 sec)
```

[illegible]innodb_rollback_on_timeout[illegible]

```
mysql[illegible]SET@@innodb_rollback_on_timeout=on;

ERROR 1238(HY000):Variable'innodb_rollback_on_timeout'is a read only variable
```

[illegible]MySQL[illegible]1205[illegible]

```
mysql[illegible]BEGIN;

Query OK,0 rows affected(0.00 sec)

mysql[illegible]SELECT*FROM t WHERE a=1 FORUPDATE;

ERROR 1205(HY000):Lock wait timeout exceeded;try restarting transaction
```

[illegible]InnoDB[illegible]InnoDB[illegible]

```
#[illegible]A

mysql[illegible]SELECT*FROM t;
```

```
+---+
|a|
+---+
|1|
|2|
|4|
+---+
3 rows in set(0.00 sec)
mysql□BEGIN;
Query OK,0 rows affected(0.00 sec)
mysql□SELECT*FROM t WHERE a□4 FOR UPDATE;
+---+
|a|
+---+
|1|
|2|
+---+
2 rows in set(0.00 sec)
```

□□□A□□□□□□□□□□Next-Key Lock□□□□□□□□4□□□□□□□□□□□□4□□□□□□□□□□□□□□B□□□□□□□□

```
#□□B
mysql□BEGIN;
Query OK,0 rows affected(0.00 sec)
mysql□INSERT INTO t SELECT 5;
Query OK,1 row affected(0.00 sec)
Records:1 Duplicates:0 Warnings:0
mysql□INSERTINTO t SELECT 3;
ERROR 1205(HY000):Lock wait timeout exceeded;try restarting transaction
```

□□□□□□□□B□□□□□5□□□□□□□□□□□□3□□□□□□A□Next-Key Lock□□
□□□□□□□□□□A□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□SELECT□
□□□□□□5□□□□□□□□□

```
mysql□SELECT*FROM t;

+---+

|a|

+---+

|1|

|2|

|4|

|5|

|8|

+---+

5 rows in set(0.00 sec)
```

□□□□□□□□B□□□□□□□□□□□□□□□□□□□COMMIT□□□□□□□□
ROLLBACK□□□□□□□□□□□□□□□□□□□□□□□□COMMIT□□
ROLLBACK□□□□□□□□□□□□□

6.7 □□

6.7.1 □□□□□

□□□

□□□InnoDB□□□□□□□□innodb_lock_wait_timeout□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□FIFO□□□undo log□□□□□FIFO□□

□□□□□□□□□□□□□□□□□□□□□wait-for graph□□InnoDB□□□□□□□□□□□□□wait-for graph□□□□□□□□□□□□□□

❑□□□□□□

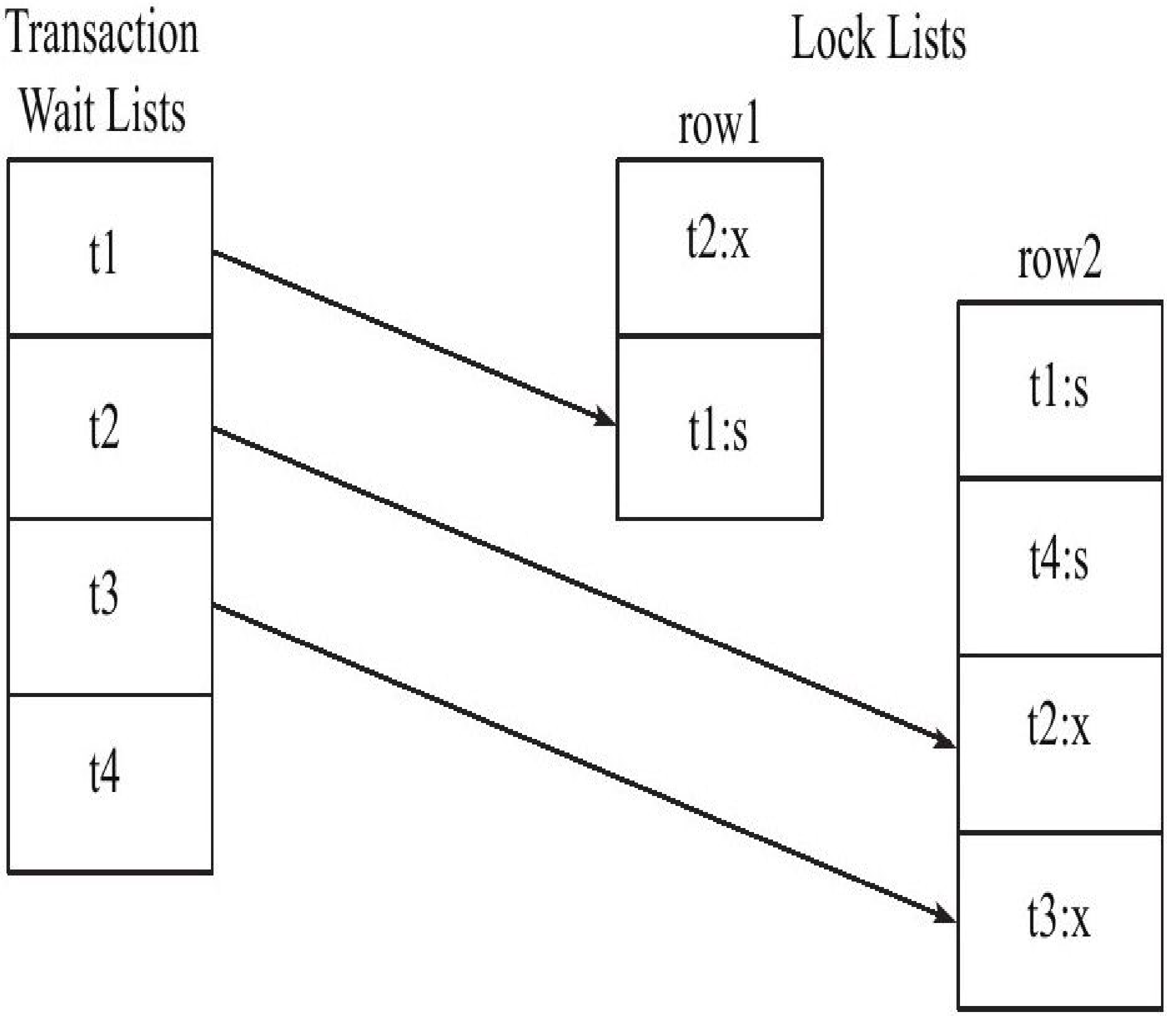
❑□□□□□□

□□□wait-for graph□□□□□□□□□□□□□□□□□□T1□□T2□□□□□□

❑□□T1□□□□T2□□□□□□

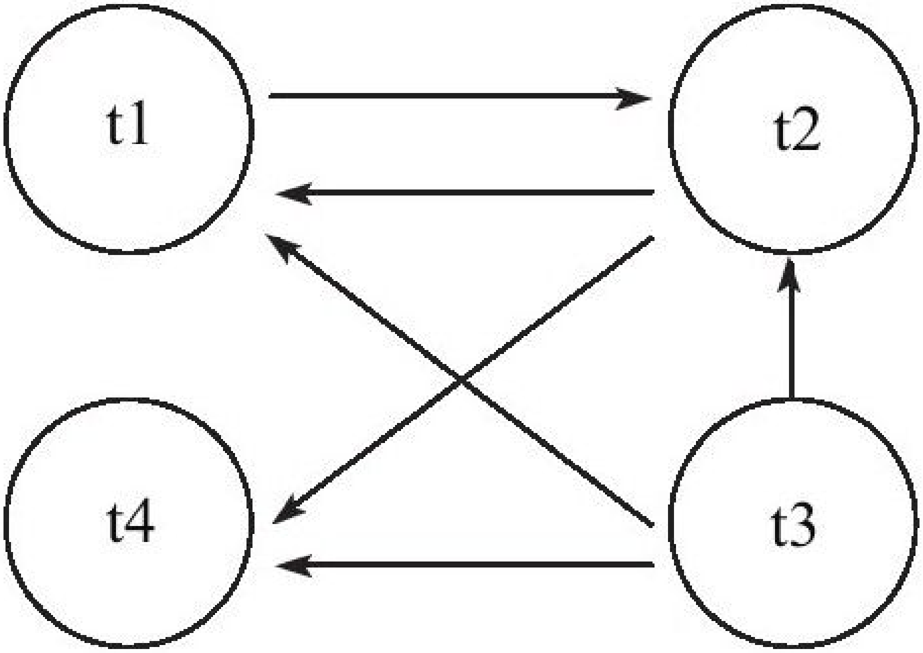
❑□□T1□□□□T2□□□□□□□□□□□□□□□□□□□□□□□□□□T1□□□□□T2□□□

□□□□□□□□□□□□□□□□□□□□6-5□□□



□ 6-5 □□□□□□□□□□

□Transaction Wait Lists□□□□□□□4□□□t1□t2□t3□t4□□□wait-for graph□□□4□□□□□□□t2□row1□□x□□□□t1□row2□□s□□□□t1□□□□□□t2□row1□□□□□□□wait-for graph□□□□□□□t1□□□□t2□□□t2□□□□□□t1□t4□□□□row2□□□□□□□□□t2□□□t1□t4□□□□□□□□□□t3□□□t1□t2□t4□□□□□□□□wait-for graph□□6-6□□□



□ 6-6 wait-for graph

□□□6-6□□□□□□□□□t1□t2□□□□□□□□□□□□□□□□□□□□□wait-for graph□□InnoDB□□□□□□□□undo□□□□□□□

wait-for graph□□□□□□□□□□□□□□□□□□□□InnoDB1.2□□□□□□□□□□□□□□□□□□1.2□□□□□□wait-for graph□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□

6.7.2 [illegible]

[illegible]2[illegible]

[illegible]n+1[illegible]n+1[illegible]r+1[illegible]R[illegible]nr[illegible]R[illegible]

[illegible]

$$(1+2+3+\cdots+r)/(r+1)\approx r/2$$

[illegible]R[illegible]1/R[illegible]PW[illegible]

$$PW=nr/2R$$

[illegible]r[illegible]PW(T)[illegible]

$$PW(T)=1-(1-PW)^{r}\approx r*PW\approx\frac{nr^{2}}{2R}$$

[illegible]2[illegible]

$$\text{一个事务发生死锁的概率}\approx\frac{PW(T)^{2}}{n}\approx\frac{nr^{4}}{4R^{2}}$$

[illegible]2[illegible]

$$\text{系统中任何一个事务发生死锁的概率} \approx \frac{n^2 r^4}{4R^2}$$

[illegible]nr[illegible]R[illegible]

- [illegible]n[illegible]
- [illegible]r[illegible]
- [illegible]R[illegible]

6.7.3 [illegible]

[illegible] 6-18 [illegible] A [illegible] B [illegible] B [illegible] A [illegible] AB-BA [illegible]

表 6-18　死锁用例 1

时间	会话 A	会话 B
1	BEGIN;	
2	mysql>SELECT * FROM t WHERE a = 1 FOR UPDATE; ********* 1. row ************ a: 1 1 row in set (0.00 sec)	BEGIN
3		mysql>SELECT * FROM t WHERE a = 2 FOR UPDATE; ********* 1. row ************ a: 2 1 row in set (0.00 sec)
4	mysql>SELECT * FROM t WHERE a = 2 FOR UPDATE; # 等待	
5		mysql>SELECT * FROM t WHERE a = 1 FOR UPDATE; ERROR 1213 (40001): Deadlock found when trying to get lock; try restarting transaction

□□□□□□□□□B□□□□□□□1213□□□□□□□□□□□□□□□□□□□□□□□□□□A□B□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□B□□□□□□□□□□□□□□A□□□□□□□□□2□□□□□□□□□□□□□□B□□□□□□□□□□□□□□A□□□□□□□□□□□□□□□□□□6.6□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□1213□□□□□□□□□□□□□□□□□□

Oracle□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□MySQL□□□□□□□□□□□

```
mysql□CREATE TABLE p(

-□aINT,

-□PRIMARY KEY(a)

-□)ENGINE=InnoDB;

Query OK,0 rows affected(0.00 sec)

mysql□CREATE TABLE c(

-□bINT,

-□FOREIGH KEY(b)REFERENCES p(a)

-□)ENGINE=InnoDB;

Query OK,0 rows affected(0.00 sec)

mysql□SHOW INDEX FROM c\G;

***************************1.row***************************

Table:c

Non_unique:1

Key_name:b

Seq_in_index:1

Column_name:b

Collation:A

Cardinality:0

Sub_part:NULL

Packed:NULL
```

```
Null:YES

Index_type:BTREE

Comment:

1 row in set(0.00 sec)

mysql[illegible]DROP INDEX b ON c;

ERROR 1553(HY000):Cannot drop index'b':needed in a foreign key constraint
```

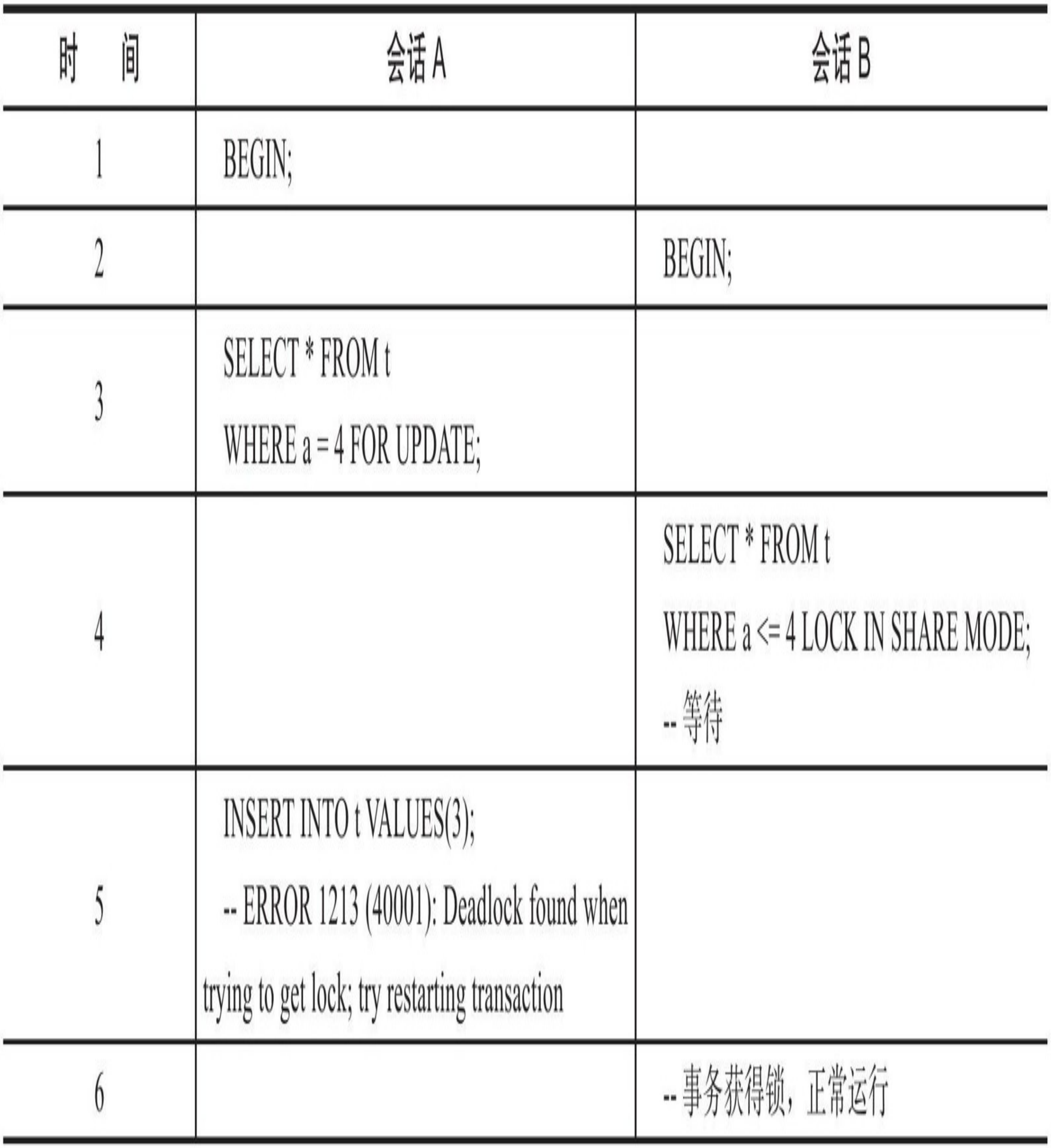
[illegible]InnoDB[illegible]b[illegible]

[illegible]X[illegible]S[illegible]t[illegible]

```
CREATE TABLE t(

a INT PRIMARY KEY

)ENGINE=InnoDB;

INSERT INTO t VALUES(1),(2),(4),(5);
```

[illegible]t[illegible]a[illegible]4[illegible]6-19[illegible]

表6-19 死锁用例2



时间	会话A	会话B
1	BEGIN;	
2		BEGIN;

3	SELECT * FROM t
	WHERE a = 4 FOR UPDATE;	

4		SELECT * FROM t
		WHERE a <= 4 LOCK IN SHARE MODE;
		-- 等待

5	INSERT INTO t VALUES(3); -- ERROR 1213 (40001): Deadlock found when trying to get lock; try restarting transaction	
6		-- 事务获得锁，正常运行

[illegible]A[illegible]4[illegible]X[illegible]A[illegible]3[illegible]B[illegible]4[illegible]S[illegible]1[illegible]2[illegible]5[illegible]B[illegible]4[illegible]S[illegible]3[illegible]

InnoDB[illegible]undo log[illegible]AB-BA[illegible]

6.8 [illegible]

[illegible]Lock Escalation[illegible]1000[illegible]

Microsoft SQL Server[illegible]

[illegible]Microsoft SQL Server 2005[illegible]SQL Server[illegible]InnoDB[illegible]

❑[illegible]SQL[illegible]5000[illegible]

❑[illegible]40%[illegible]

[illegible]Microsoft SQL Server[illegible]

InnoDB[illegible]

[illegible]3 000 000[illegible]100[illegible]300 000 000[illegible]SQL[illegible]X[illegible]10[illegible]3GB[illegible]InnoDB[illegible]30[illegible]90MB[illegible]

6.9 [illegible]

[illegible] MySQL [illegible]

[illegible] MySQL [illegible] InnoDB [illegible] MyISAM [illegible] Microsoft SQL Server [illegible] Oracle [illegible] SQL [illegible] InnoDB [illegible]

[illegible]7[illegible] [illegible]

[illegible]Transaction[illegible]

[illegible]

InnoDB[illegible]ACID[illegible]ACID[illegible]4[illegible]

❏[illegible]atomicity[illegible]

❏[illegible]consistency[illegible]

❏[illegible]isolation[illegible]

❏[illegible]durability[illegible]

[illegible]6[illegible]InnoDB[illegible]

7.1 [illegible]

7.1.1 [illegible]

[illegible]SQL[illegible]SQL[illegible]

[illegible]ACID[illegible]ACID[illegible]MySQL[illegible]NDB Cluster[illegible]D[illegible]Oracle[illegible]READ COMMITTED[illegible]I[illegible]InnoDB[illegible]READ REPEATABLE[illegible]ACID[illegible]ACID[illegible]

A[illegible]Atomicity[illegible]C[illegible]SQRT[illegible]SQRT[illegible]

[illegible]ATM[illegible]

1[illegible]ATM[illegible]

2[illegible]

3[illegible]ATM[illegible]

4[illegible]

5[illegible]ATM[illegible]

6[illegible]

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ATM□□□□□□□□□□□ □□□

□□ □□□□□□□□□□SQL□□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□□ □□□□□□

□□□ □□□ □□□ □□□□□□□□□□□□□□□□□□

C□consistency□□ □□□ □□□ □□□ □□□□□□□□□□□□□□□□□□□□□□□□□□□□□——□□□□□□□□□

I□isolation□□□□□□□□□□□□□□□□□□□□□□□□concurrency control□□ □□□□□serializability□□□□locking□□□□□□□□□□□□□□□□□□□□□□□□□ □□□ □□□□□□□granular lock□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ □□□□

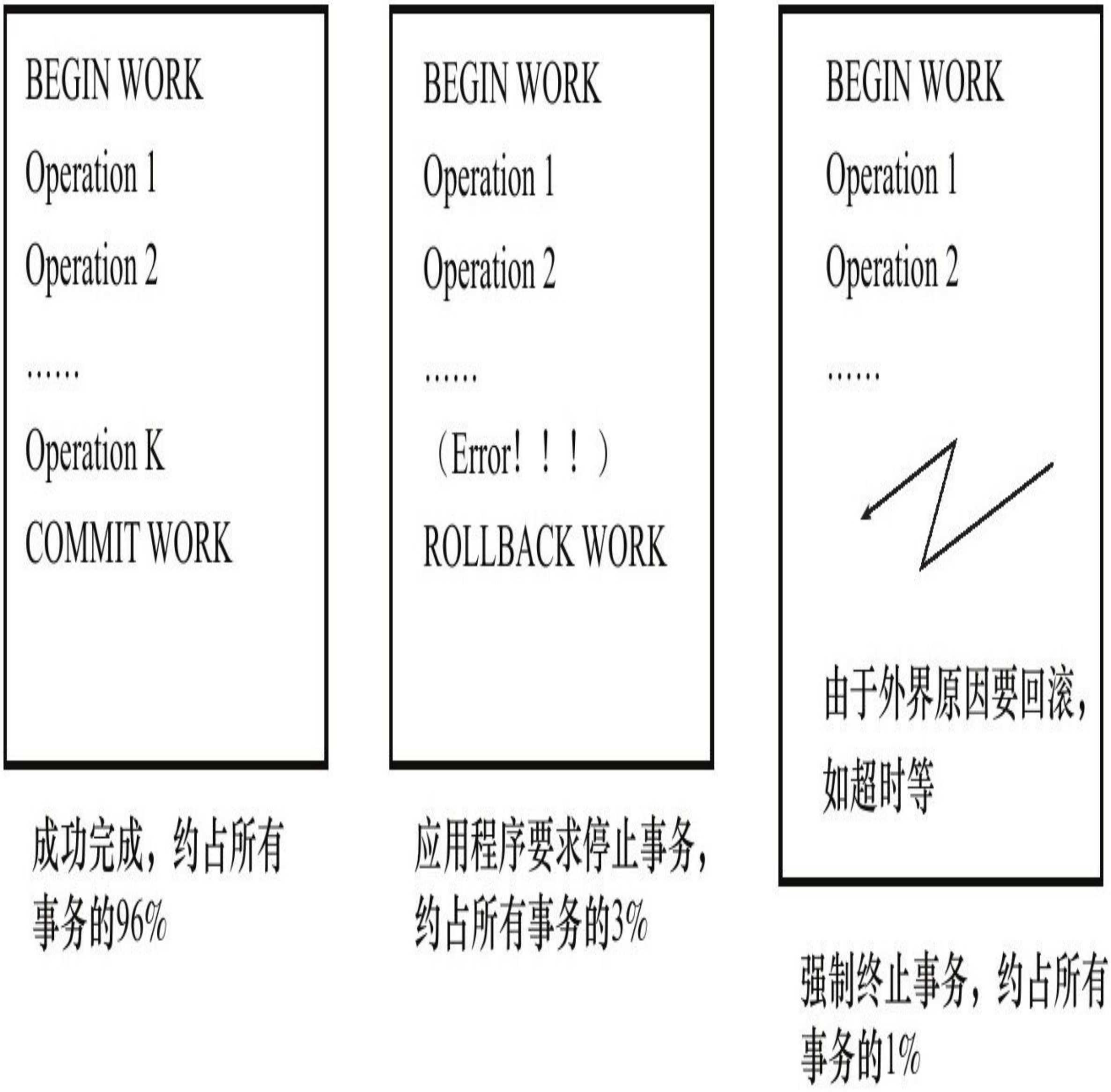
D□durability□□ □□□ □□ □□□□□□□□□□□□□□RAID□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□□□High Reliability□□□□□□□□□□□High Availability□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

7.1.2 [illegible]

[illegible]

- ❑[illegible]Flat Transactions[illegible]
- ❑[illegible]Flat Transactions with Savepoints[illegible]
- ❑[illegible]Chained Transactions[illegible]
- ❑[illegible]Nested Transactions[illegible]
- ❑[illegible]Distributed Transactions[illegible]

[illegible]Flat Transaction[illegible]BEGIN WORK[illegible]COMMIT WORK[illegible]ROLLBACK WORK[illegible]7-1[illegible]


BEGIN WORK
Operation 1
Operation 2
……
Operation K
COMMIT WORK
成功完成，约占所有事务的96%
BEGIN WORK
Operation 1
Operation 2
……
（Error！！！）
ROLLBACK WORK
应用程序要求停止事务，约占所有事务的3%
BEGIN WORK
Operation 1
Operation 2
……
由于外界原因要回滚，如超时等
强制终止事务，约占所有事务的1%

[illegible] 7-1 [illegible]

[illegible]7-1[illegible]

[illegible]

BEGIN WORK

S1[illegible]

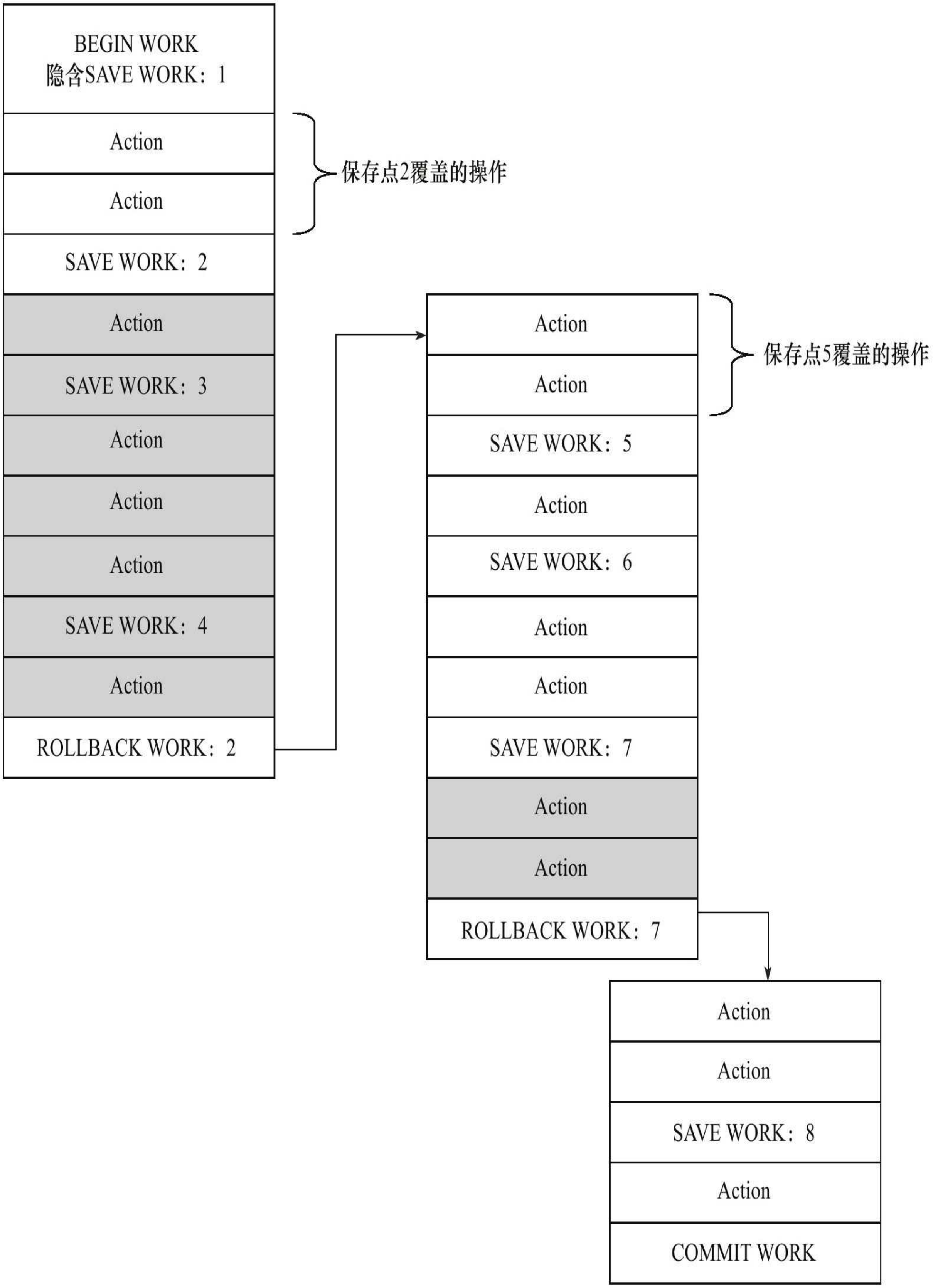
S2[illegible]

S3[illegible]

[illegible]S3[illegible]S1[illegible]S2[illegible]S3[illegible]S1[illegible]S2[illegible]

[illegible]Flat Transactions with Savepoint[illegible]Savepoint[illegible]

[illegible]SAVE WORK[illegible]7-2[illegible]


BEGIN WORK
隐含SAVE WORK：1
Action
Action
保存点2覆盖的操作
SAVE WORK：2
Action
SAVE WORK：3
Action
Action
Action
SAVE WORK：4
Action
ROLLBACK WORK：2
Action
Action
保存点5覆盖的操作
SAVE WORK：5
Action
SAVE WORK：6
Action
Action
SAVE WORK：7
Action
Action
ROLLBACK WORK：7
Action
Action
SAVE WORK：8
Action
COMMIT WORK

[illegible] 7-2 [illegible]

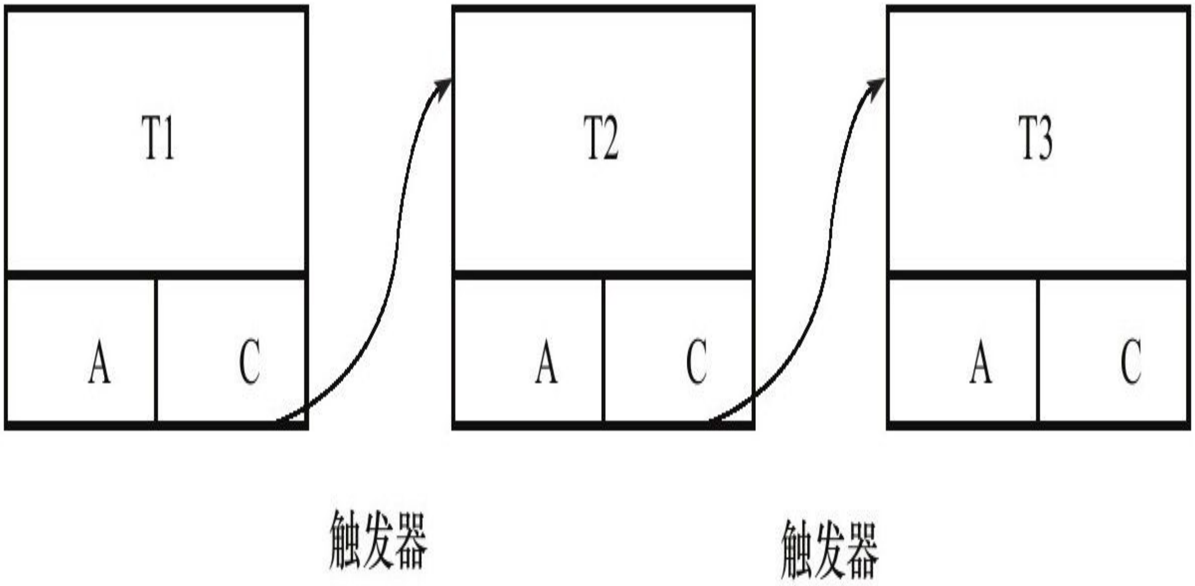
[illegible]7-2[illegible]ROLLBACK WORK[illegible]BEGIN WORK[illegible]ROLLBACK WORK:2[illegible]2[illegible]ROLLBACK WORK:7[illegible]COMMIT WORK[illegible]

[illegible]7-2[illegible]2[illegible]3[illegible]5[illegible]ROLLBACK[illegible]

[illegible]ROLLBACK WORK:2[illegible]2[illegible]ROLLBACK WORK[illegible]

[illegible]Chained Transaction[illegible]volatile[illegible]persistent[illegible]

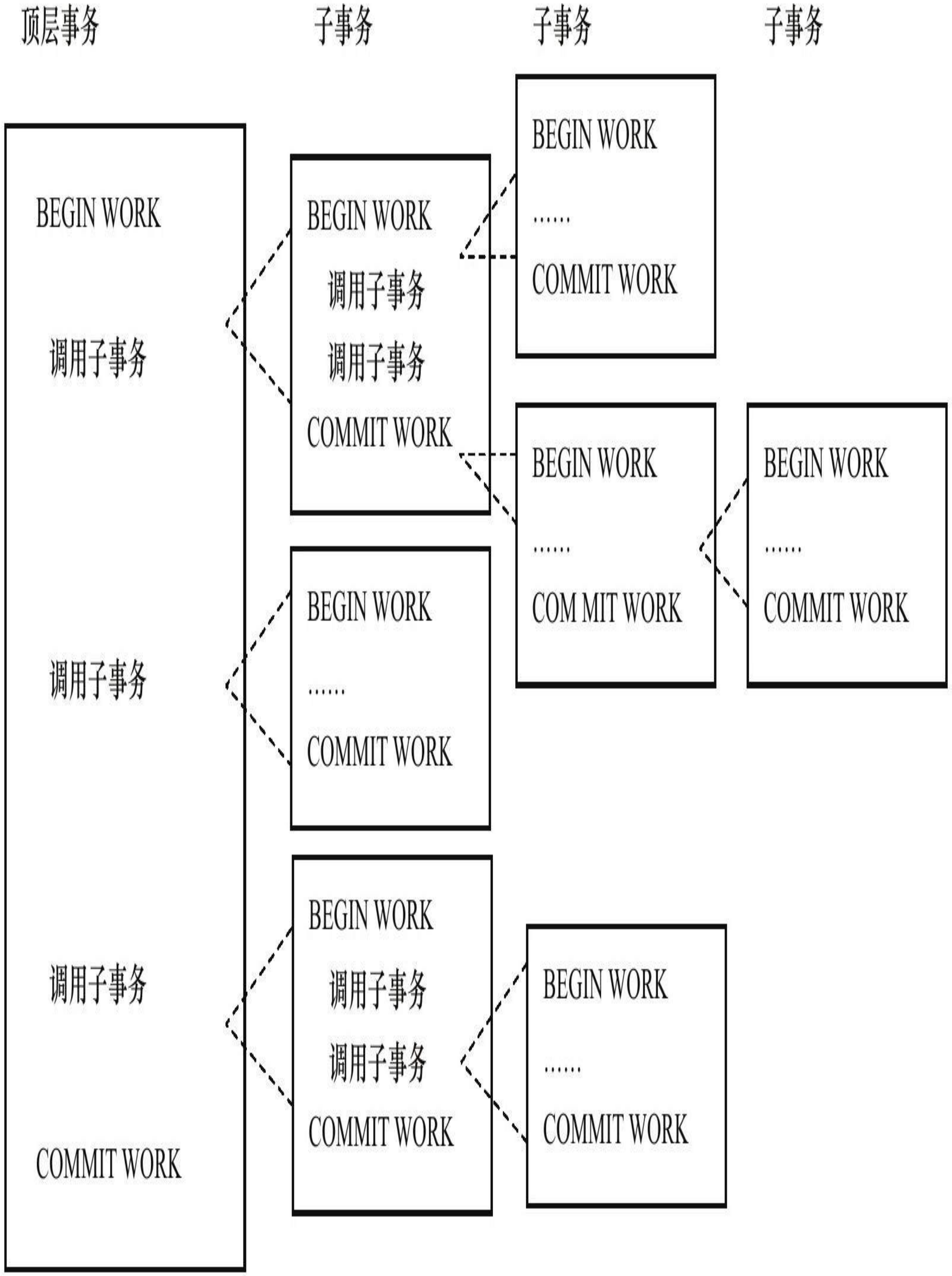
[illegible]7-3[illegible]



[illegible] 7-3 [illegible]

[illegible]COMMIT[illegible]

[illegible]Nested Transaction[illegible]top-level transaction[illegible]subtransaction[illegible]7-4[illegible]


顶层事务
子事务
子事务
子事务
BEGIN WORK
调用子事务
调用子事务
调用子事务
COMMIT WORK
BEGIN WORK
调用子事务
调用子事务
COMMIT WORK
BEGIN WORK
……
COMMIT WORK
BEGIN WORK
……
COM MIT WORK
BEGIN WORK
……
COMMIT WORK
BEGIN WORK
……
COMMIT WORK
BEGIN WORK
调用子事务
调用子事务
COMMIT WORK
BEGIN WORK
……
COMMIT WORK

[illegible] 7-4 [illegible]

[illegible]Moss[illegible]

1[illegible]

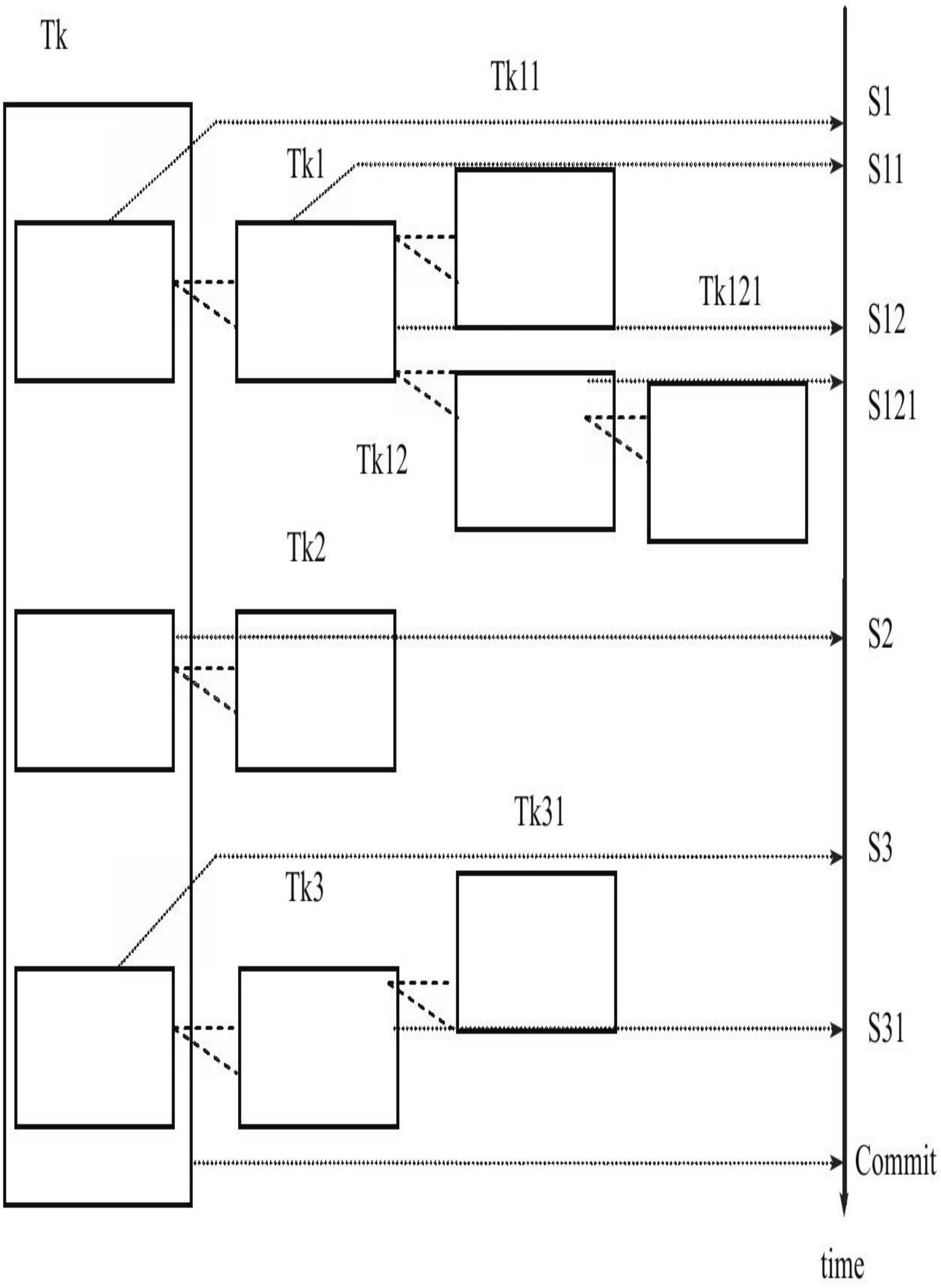
2[illegible]

3[illegible]predecessor[illegible]parent[illegible]child[illegible]

4[illegible]

5[illegible]A[illegible]C[illegible]I[illegible]D[illegible]

[illegible]Moss[illegible]7-5[illegible]


Tk
Tk11
S1
Tk1
S11
Tk121
S12
S121
Tk12
Tk2
S2
Tk31
S3
Tk3
S31
Commit
time

[illegible] 7-5 [illegible]

[illegible]7-5[illegible]Tk3[illegible]S2[illegible]

[illegible]P_1[illegible]X[illegible]Y[illegible]P_{11}[illegible]P_1[illegible]P_{11}[illegible]Z[illegible]counter-inherited[illegible]P_1[illegible]3[illegible]X[illegible]Y[illegible]Z[illegible]P_{12}[illegible]

[illegible]

[illegible]Distributed Transactions[illegible]

[illegible]ATM[illegible]10 000[illegible]ATM[illegible]A[illegible]B[illegible]C[illegible]

1[illegible]A[illegible]

2[illegible]B[illegible]10 000[illegible]

3[illegible]C[illegible]10 000[illegible]

4[illegible]A[illegible]A[illegible]

[illegible]A[illegible]ACID[illegible]2[illegible]3[illegible]

[illegible]InnoDB[illegible]MySQL[illegible]InnoDB[illegible]

7.2 □□□□□

□□□□□□□6□□□□□□□□□□□□□□□□□□□□□□□□□□redo log□undo log□□□□redo log□□□□□□□□□□□□□□□□□□□□□□undo log□□□□□□□□□□□

□□DBA□□□□□undo□redo□□□□□□□□□□redo□undo□□□□□□□□□□□□□□□□redo□□□□□□□□□□□□□□undo□□□□□□□□□□□□□□□□□□□□□□□□□redo□□□□□□□□□□□□□□□□□□□□□undo□□□□□□□□□□□□□□□□□

7.2.1 redo

1.□□□□

□□□□□□□□□□□□□□□□□□ACID□□D□□□□□□□□□□□□□□□□□□□□□□redo log buffer□□□□□□□□□□□□□□□□□redo log file□□□□□□□□

InnoDB□□□□□□□□□□□□Force Log at Commit□□□□□□□□□□□□□□□□□□COMMIT□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□COMMIT□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□redo log□undo log□redo log□□□□□□□□□□□undo log□□□□□□□□□MVCC□□□□redo log□□□□□□□□□□□□□□□□□□□□□redo log□□□□□□□□□□□undo log□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□fsync□□□□□□□□□□□□□□□□□□O_DIRECT□□fsync□□□□□fsync□□

InnoDB□□□fsync□□□□□□□□□□□□□□□□□□□fsync□□□

□□innodb_flush_log_at_trx_commit□□□□□□□□□□□□□□□□□□□□□□□□□1□□□□□□□□□□□□□□fsync□□□□□□□□□□□□□□0□2□0□□□□□□□□□□□□□□□□□□□□□□□□□master thread□□□□□□master thread□□1□□

[illegible]fsync[illegible]2[illegible]fsync[illegible]MySQL[illegible]

[illegible]innodb_flush_log_at_trx_commit[illegible]t1[illegible]p_load[illegible]

```
CREATE TABLE test_load(

a INT,

b CHAR(80)

)ENGINE=INNODB;

DELIMITER//

CREATE PROCEDURE p_load(count INT UNSIGNED)

BEGIN

DECLARE s INT UNSIGNED DEFAULT 1;

DECLARE c CHAR(80)DEFAULT REPEAT('a',80);

WHILE s[illegible]=count DO

INSERT INTO test_load SELECT NULL,c;

COMMIT;

SET s=s+1;

END WHILE;

END;

//

DELIMITER;
```

[illegible]p_load[illegible]test_load[illegible]COMMIT[illegible]innodb_flush_log_at_trx_commit[illegible]1[illegible]InnoDB[illegible]fsync[illegible]CALL p_load[illegible]500 000[illegible]50[illegible]50[illegible]fsync[illegible]50[illegible]

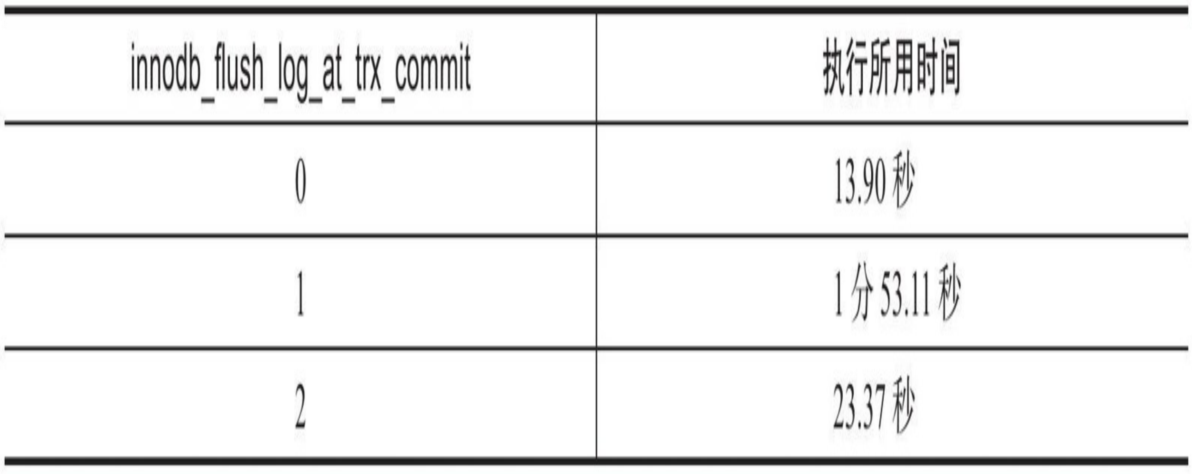
```
mysql□CALL p_load(500000);

Query OK,0 rows affected(1 min 53.11 sec)
```

[illegible]50[illegible]2[illegible]fsync[illegible] innodb_flush_log_at_trx_commit[illegible]0[illegible]

```
mysql□SHOW VARIABLES LIKE'innodb_flush_log_at_trx_commit'\G

***************************1.row***************************

Variable_name:innodb_flush_log_at_trx_commit

Value:0

1 row in set(0.00 sec)

mysql□CALL p_load(500000);

Query OK,0 rows affected(13.90 sec)
```

[illegible]innodb_flush_log_at_trx_commit[illegible]0[illegible]50[illegible]13.90[illegible]12%[illegible]fsync[illegible]7-1[illegible] innodb_flush_log_at_trx_commit[illegible]p_load[illegible]50[illegible]

表 7-1 不同 innodb_flush_log_at_trx_commit 设置对于插入的速度影响



innodb_flush_log_at_trx_commit	执行所用时间
0	13.90 秒
1	1 分 53.11 秒
2	23.37 秒

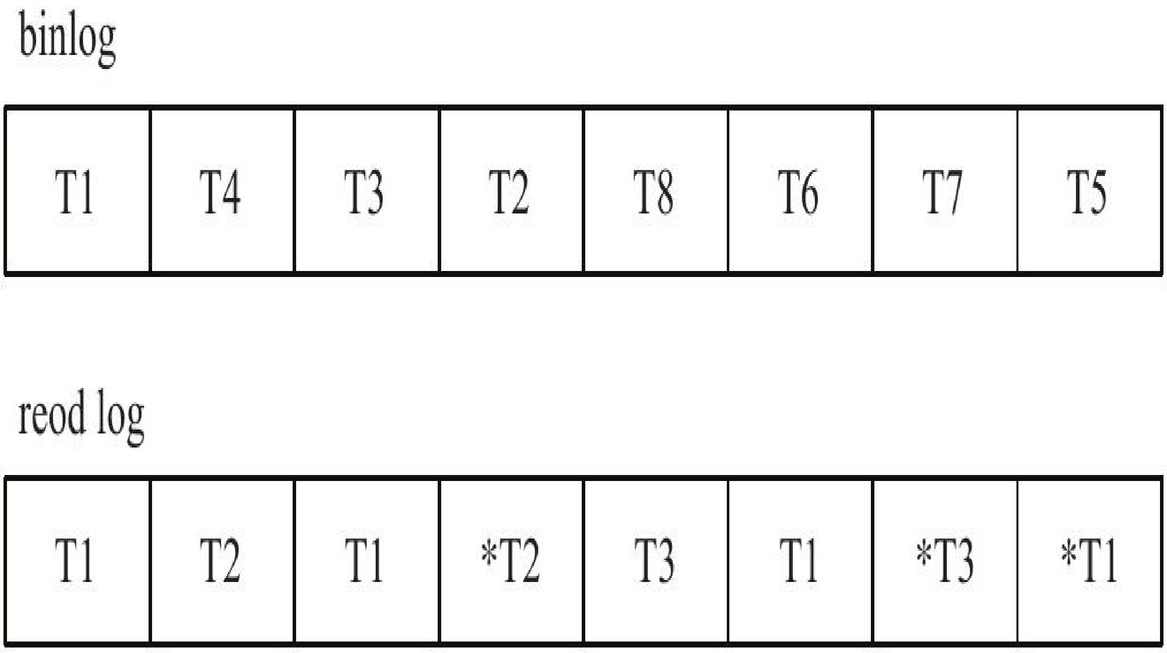
[illegible]innodb_flush_log_at_trx_commit[illegible]0[illegible]2[illegible]ACID[illegible]50[illegible]COMMIT[illegible]COMMIT[illegible]

[illegible]MySQL[illegible]binlog[illegible]POINT-IN-TIME[illegible]PIT[illegible]Replication[illegible]

[illegible]InnoDB[illegible]MySQL[illegible]InnoDB[illegible]MySQL[illegible]

[illegible]MySQL[illegible]SQL[illegible]InnoDB[illegible]

[illegible]7-6[illegible]InnoDB[illegible]



[illegible] 7-6 [illegible]

[illegible]7-6[illegible]InnoDB[illegible]*T1[illegible]*T2[illegible]*T3[illegible]

2.log block

[illegible]InnoDB[illegible]512[illegible]block[illegible]redo log block[illegible]512[illegible]

[illegible]512[illegible]512[illegible]doublewrite[illegible]

[illegible]log block header[illegible]log block tailer[illegible]12[illegible]8[illegible]492[illegible]512-12-8[illegible]7-7[illegible]

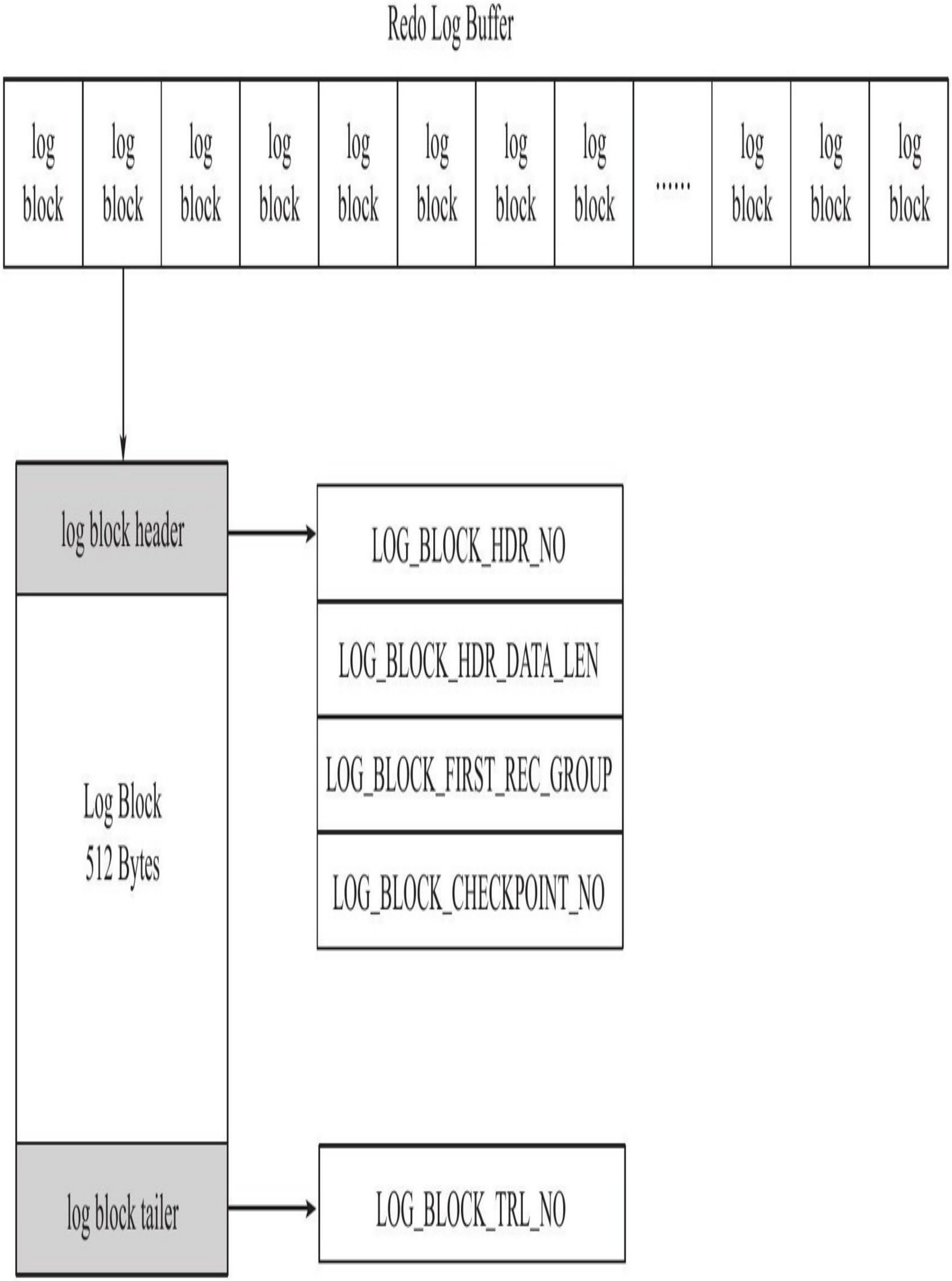
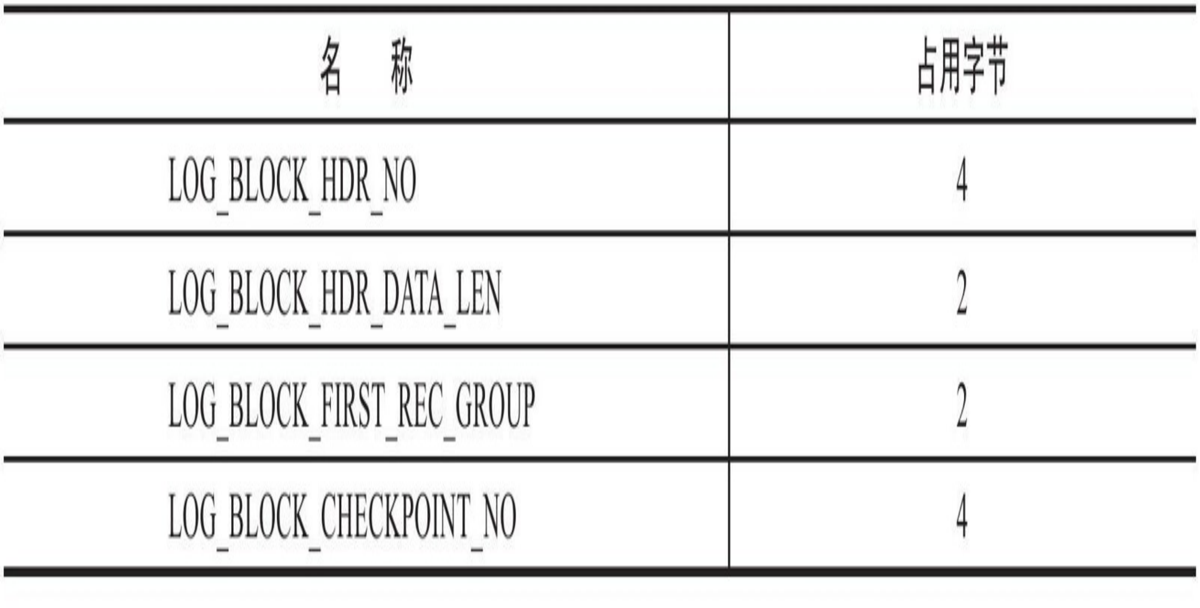

Redo Log Buffer
log block
log block
log block
log block
log block
log block
log block
log block
……
log block
log block
log block
log block header
LOG_BLOCK_HDR_NO
LOG_BLOCK_HDR_DATA_LEN
LOG_BLOCK_FIRST_REC_GROUP
LOG_BLOCK_CHECKPOINT_NO
Log Block
512 Bytes
log block tailer
LOG_BLOCK_TRL_NO

图 7-7 □□□□□□□□□□

□7-7□□□□□□□□□□□□□□□□□□□□□□□□□□□512□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□log block header□□□□□□□log body□□□□□□□log block tailer□□

log block header□4□□□□□□□7-2□□□

表 7-2 log block header

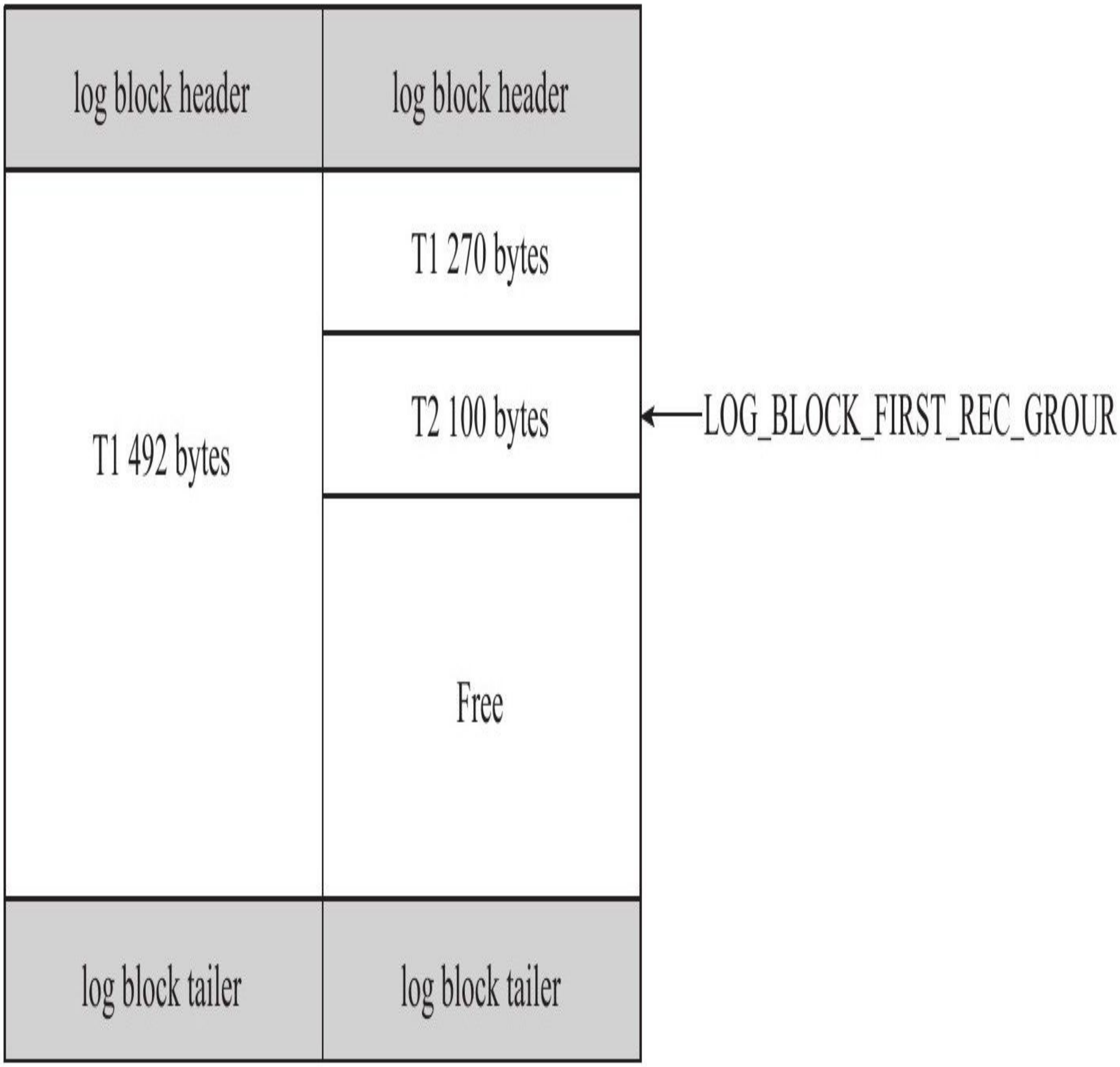


名 称	占用字节
LOG_BLOCK_HDR_NO	4
LOG_BLOCK_HDR_DATA_LEN	2
LOG_BLOCK_FIRST_REC_GROUP	2
LOG_BLOCK_CHECKPOINT_NO	4

log buffer□□log block□□□□□□log buffer□□□□□□□□□□LOG_BLOCK_HDR_NO□□□□□□□□□□□□□□□□□□□□□□□□□□□4□□□□□□□□□□□□□□□□□□flush bit□□□□□□□□2G□

LOG_BLOCK_HDR_DATA_LEN□□2□□□□□log block□□□□□□□□log block□□□□□□□□0x200□□□□□□□log block□□□□□□512□□□

LOG_BLOCK_FIRST_REC_GROUP□□2□□□□□□log block□□□□□□□□□□□□□□□□□□□□□LOG_BLOCK_HDR_DATA_LEN□□□□□□□□log block□□□□□□□□□□□T1□□□□□1□□762□□□□□T2□□□□□□100□□□□□□□log block□□□□□□492□□□□□□□□log buffer□□□□□□□7-8□□□



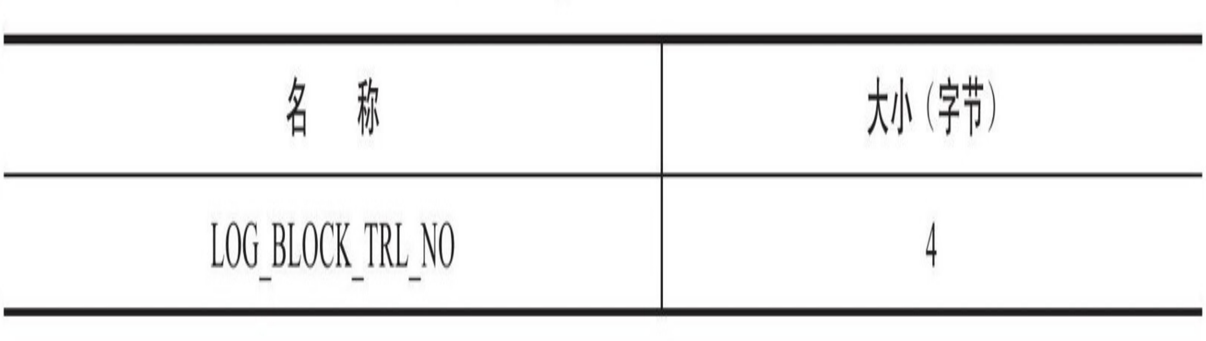
□ 7-8 LOG_BLOCK_FIRST_REC_GROUP□□□

□□7-8□□□□□□□□□□□T1□□□□□□□792□□□□□□□□□□□log block□□□□log block□LOG_BLOCK_FIRST_REC_GROUP□12□□log block□□□□□□□□□□□□□□□□log block□□□□□□□□□□□□□T1□□□□□□□□T2□□□□□log block□□□□□□□□□□log block□LOG_BLOCK_FIRST_REC_GROUP□282□270+12□□

LOG_BLOCK_CHECKPOINT_NO□□4□□□□□□log block□□□□□□□□□□□4□□□□□

log block tailer[illegible]1[illegible]7-3[illegible]LOG_BLOCK_HDR_NO[illegible]log_block_init[illegible]

表7-3 log block tailer部分



名称	大小（字节）
LOG_BLOCK_TRL_NO	4

3.log group

log group[illegible]log group[illegible]ha_innobase.cc[illegible]InnoDB[illegible]log group[illegible]

log group[illegible]log group[illegible]log group[illegible]log group[illegible]InnoDB 1.2[illegible]4GB[illegible]4GB[illegible]InnoDB 1.2[illegible]512GB[illegible]InnoSQL[illegible]InnoDB[illegible]1.1[illegible]4GB[illegible]

[illegible]log buffer[illegible]log block[illegible]log block[illegible]512[illegible]InnoDB[illegible]log buffer[illegible]log block[illegible]

❑[illegible]

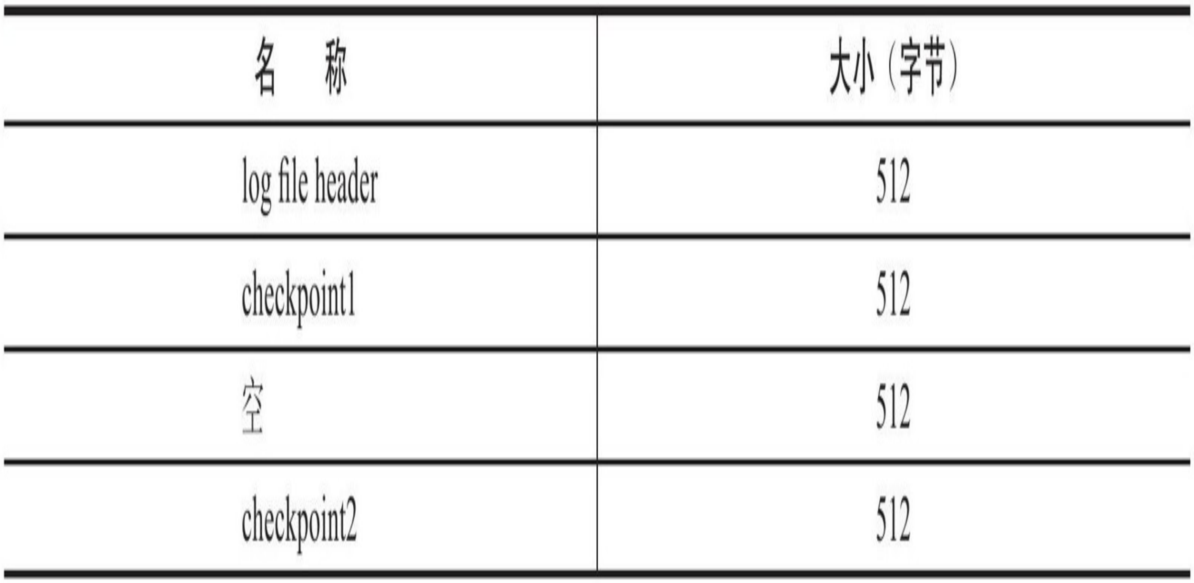
❑[illegible]log buffer[illegible]

❑log checkpoint[illegible]

[illegible]log block[illegible]append[illegible]redo log file[illegible]redo log file[illegible]redo log file[illegible]round-robin[illegible]

□□log block□□□redo log file□□□□□□□□□□□□□□□□□□□□redo log file □□□□□□□□□□□□□□□□□redo log file□□□□log buffer□□□□□□log block□□□□□□□□□□□□□□□□□□□□□□2KB□□□□□□redo log file□□2KB□□□□□□log block□□□□□□log group□□□□□redo log file□□□2KB□□□□□4□512□□□□□□□□□□□□□□□□□7-4□□□

表7-4 redo log file前2KB部分的内容



名　称	大小（字节）
log file header	512
checkpoint1	512
空	512
checkpoint2	512

□□□□□□□□□□□□□□□□□□log group□□□□redo log file□□□□□□log group□□□□redo log file□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□redo log file□□□□□□□□□□□□□□□□□□log block□□□□□□□□□□□□2KB □□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□log group□redo log file□□□□□□□7-9□□□

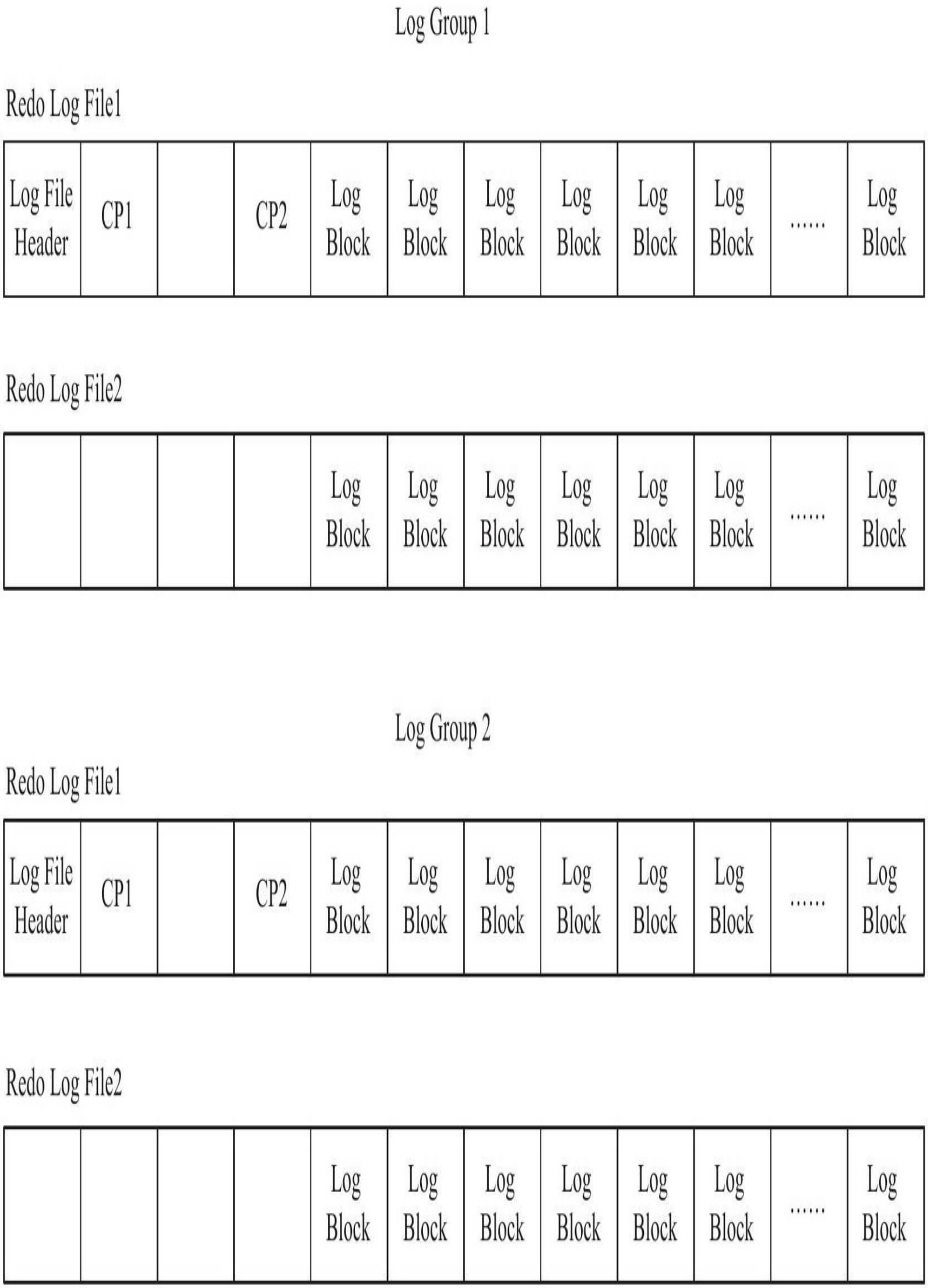

Log Group 1
Redo Log File1
Log File Header
CP1
CP2
Log Block
Log Block
Log Block
Log Block
Log Block
Log Block
......
Log Block
Redo Log File2
Log Block
Log Block
Log Block
Log Block
Log Block
Log Block
......
Log Block
Log Group 2
Redo Log File1
Log File Header
CP1
CP2
Log Block
Log Block
Log Block
Log Block
Log Block
Log Block
......
Log Block
Redo Log File2
Log Block
Log Block
Log Block
Log Block
Log Block
Log Block
......
Log Block

图 7-9 log group□redo log file□□□□□

□log filer header□□□□□□InnoDB□□□□□□□checkpoint□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□checkpoint□□□□

4.□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□
□□7-10□□□

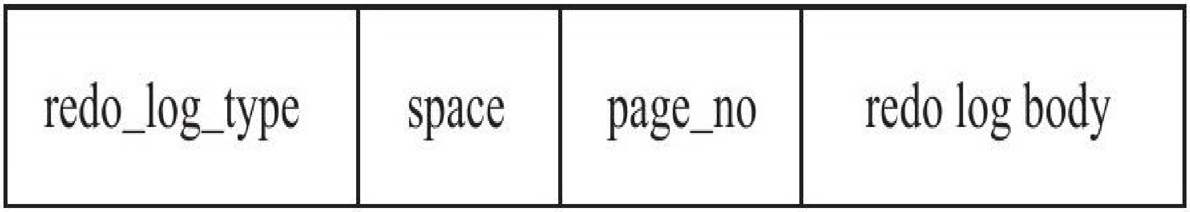
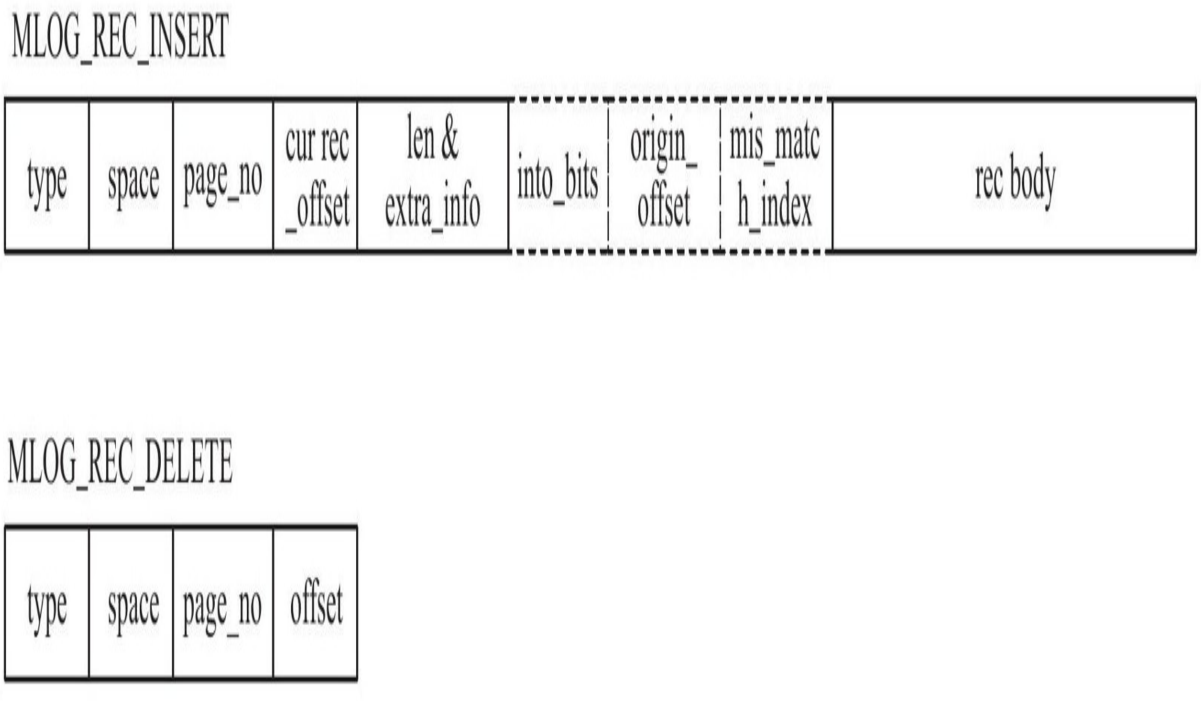


图 7-10 □□□□□□

□□□□□□□□□□3□□□□□

- redo_log_type□□□□□□□□□
- space□□□□□ID□
- page_no□□□□□□□

□□redo log body□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□7-11□□□□□□



□ 7-11 □□□□□□□□□□□□

□InnoDB1.2□□□□□□□51□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

5.LSN

LSN□Log Sequence Number□□□□□□□□□□□□□□□□InnoDB□□□□□□LSN□□8□□□□□□□□□□LSN□□□□□□□

❑□□□□□□□□□

❑checkpoint□□□

❑□□□□

LSN□□□□□□□□□□□□□□□□□□□□□□□□□□LSN□1 000□□□□□□T1□□□100□□□□□□□□□□LSN□□□□1100□□□□□□T2□□□200□□□□□□□□□□LSN□□□□1 300□□□LSN□□□□□□□□□□□□□□□□□□□

LSN□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□FIL_PAGE_LSN□□□□□□□LSN□□□□□LSN□□□□□□□□□LSN□□□□□□□□□□□□□□□□□□□□□□□□□LSN□□□□□□□□□□□□□□□□□□□P1□LSN□10 000□□□□□□□□□InnoDB□□□□□□□□□□□LSN□13 000□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□P1□□□□□□□□□□□□□□LSN□□P1□□LSN□□□□□□□□□□□P1□□□LSN□□□□□□□□□□□□□

□□□□□□□□SHOW ENGINE INNODB STATUS□□LSN□□□□

```
mysql□SHOW ENGINE INNODB STATUS\G;

……

---

LOG

---

Log sequence number 11 3047174608

Log flushed up to 11 3047174608

Last checkpoint at 11 3047174608

0 pending log writes,0 pending chkp writes

142 log i/o's done,0.00 log i/o's/second

……

1 row in set(0.00 sec)
```

Log sequence number□□□□□LSN□Log flushed up to□□□□□□□□□□□□LSN□Last checkpoint at□□□□□□□□LSN□

□□□□□□□□□□Log sequence number□Log flushed up to□□

```
mysql□show engine innodb status\G;

---

LOG
```

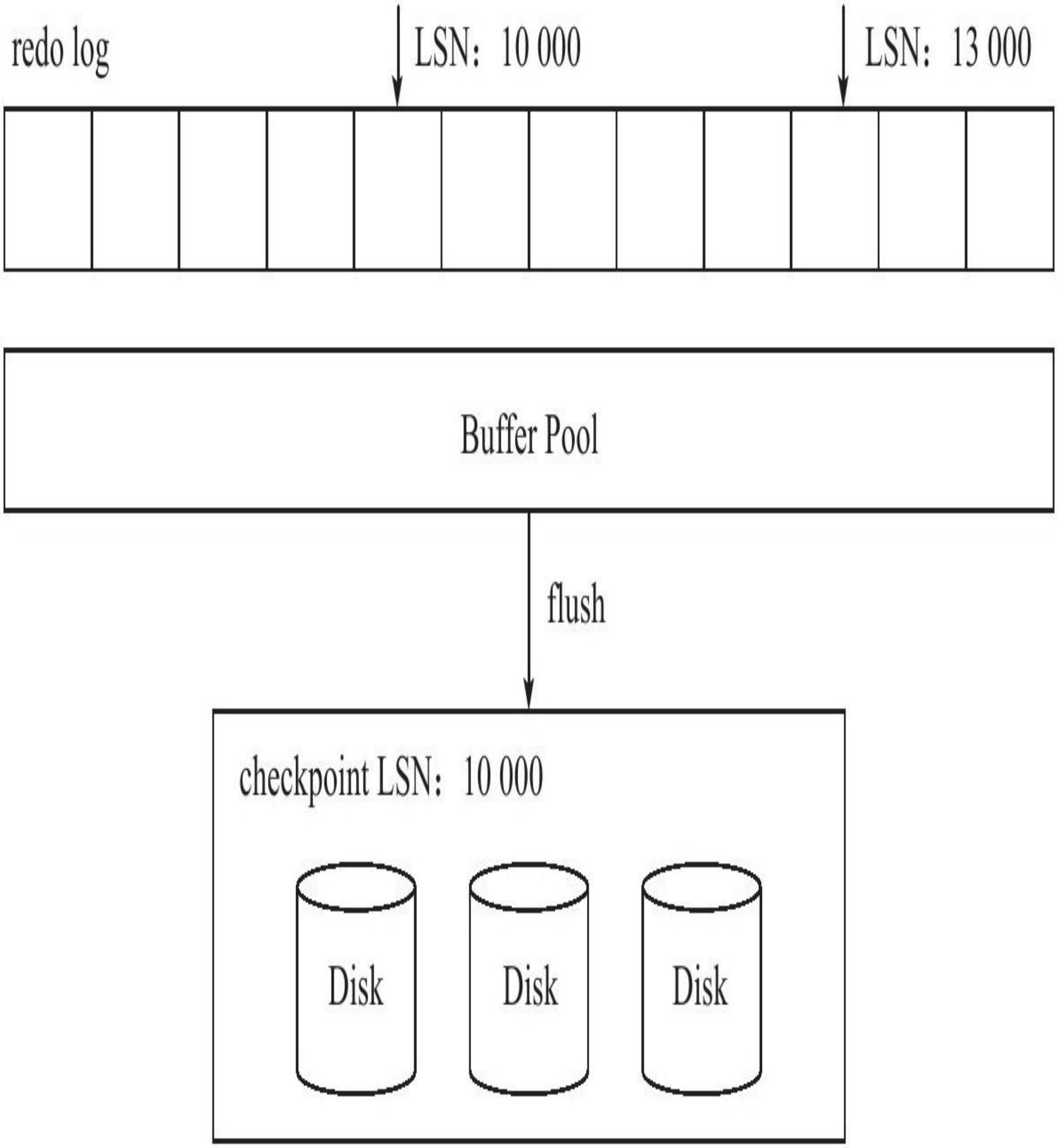
```
---

Log sequence number 203318213447

Log flushed up to 203318213326

Last checkpoint at 203252831194

1 pending log writes,0 pending chkp writes

103447 log i/o's done,7.00 log i/o's/second

……

1 row in set(0.00 sec)
```

[illegible]Log sequence number[illegible]Log flushed up to[illegible]Last checkpoint at[illegible]

6.[illegible]

InnoDB[illegible]InnoDB[illegible]

[illegible]checkpoint[illegible]LSN[illegible]checkpoint[illegible]7-12[illegible]checkpoint[illegible]LSN[illegible]10 000[illegible]LSN 10 000[illegible]13 000[illegible]



[illegible] 7-12 [illegible]

InnoDB[illegible]INSERT[illegible]

```
CREATE TABLE t(a INT,b INT,PRIMARY KEY(a),KEY(b));
```

[illegible]SQL[illegible]

```
INSERT INTO t SELECT 1,2;
```

[illegible]

```
page(2,3),offset 32,value 1,2#[illegible]

page(2,4),offset 64,value 2#[illegible]
```

[illegible]B+[illegible]split[illegible]

$$f(f(x))=f(x)$$

[illegible]DBA[illegible]ROW[illegible]INSERT[illegible]INSERT[illegible]

7.2.2 undo

1.□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□“□□”□□□□□□□□□□□□□□□□□□□□□□□□□undo□□□□□□□□□□□□□□InnoDB□□□□□□□□□redo□□□□□□□□□undo□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ROLLBACK□□□□□□□□□□□□□□undo□□□□□□□□□□□□□□□□

redo□□□□□□□□□□□□redo□□□undo□□□□□□□□□□□□□□□segment□□□□□□□□undo□□undo segment□□undo□□□□□□□□□□□□□□py_innodb_page_info.py□□□□□□□□□□□□□undo□□□□□□□□□□□□□□□□□□ibdata1□□2222□undo□□

```
[root@xen-server□]#python py_innodb_page_info.py/usr/local/mysql/data/ibdata1

Total number of page:46208:

Insert Buffer Free List:13093

Insert Buffer Bitmap:3

System Page:5

Transaction system Page:1

Freshly Allocated Page:4579

undo Log Page:2222

File Segment inode:6

B-tree Node:26296

File Space Header:1

□□□□□:2
```

□□□□□undo□□□□□□□undo□□□□□□□□□□□□□□□□□□□□□□□□——□□□□□□□□undo□□

[illegible]

[illegible]INSERT 10W[illegible]ROLLBACK[illegible]InnoDB[illegible]INSERT[illegible]InnoDB[illegible]DELETE[illegible]DELETE[illegible]InnoDB[illegible]INSERT[illegible]UPDATE[illegible]InnoDB[illegible]UPDATE[illegible]

[illegible]undo[illegible]MVCC[illegible]InnoDB[illegible]MVCC[illegible]undo[illegible]undo[illegible]

[illegible]undo log[illegible]redo log[illegible]undo log[illegible]redo log[illegible]undo log[illegible]

2.undo[illegible]

InnoDB[illegible]undo[illegible]InnoDB[illegible]rollback segment[illegible]1024[illegible]undo log segment[illegible]undo log segment[illegible]undo[illegible]5[illegible]0[illegible]5[illegible]rollback segment header[illegible]FIL_PAGE_TYPE_SYS[illegible]

[illegible]InnoDB1.1[illegible]1.1[illegible]rollback segment[illegible]1024[illegible]1.1[illegible]InnoDB[illegible]128[illegible]rollback segment[illegible]128*1024[illegible]

[illegible]InnoDB1.1[illegible]128[illegible]rollback segment[illegible]rollback segment[illegible]InnoDB1.2[illegible]rollback segment[illegible]

❑innodb_undo_directory

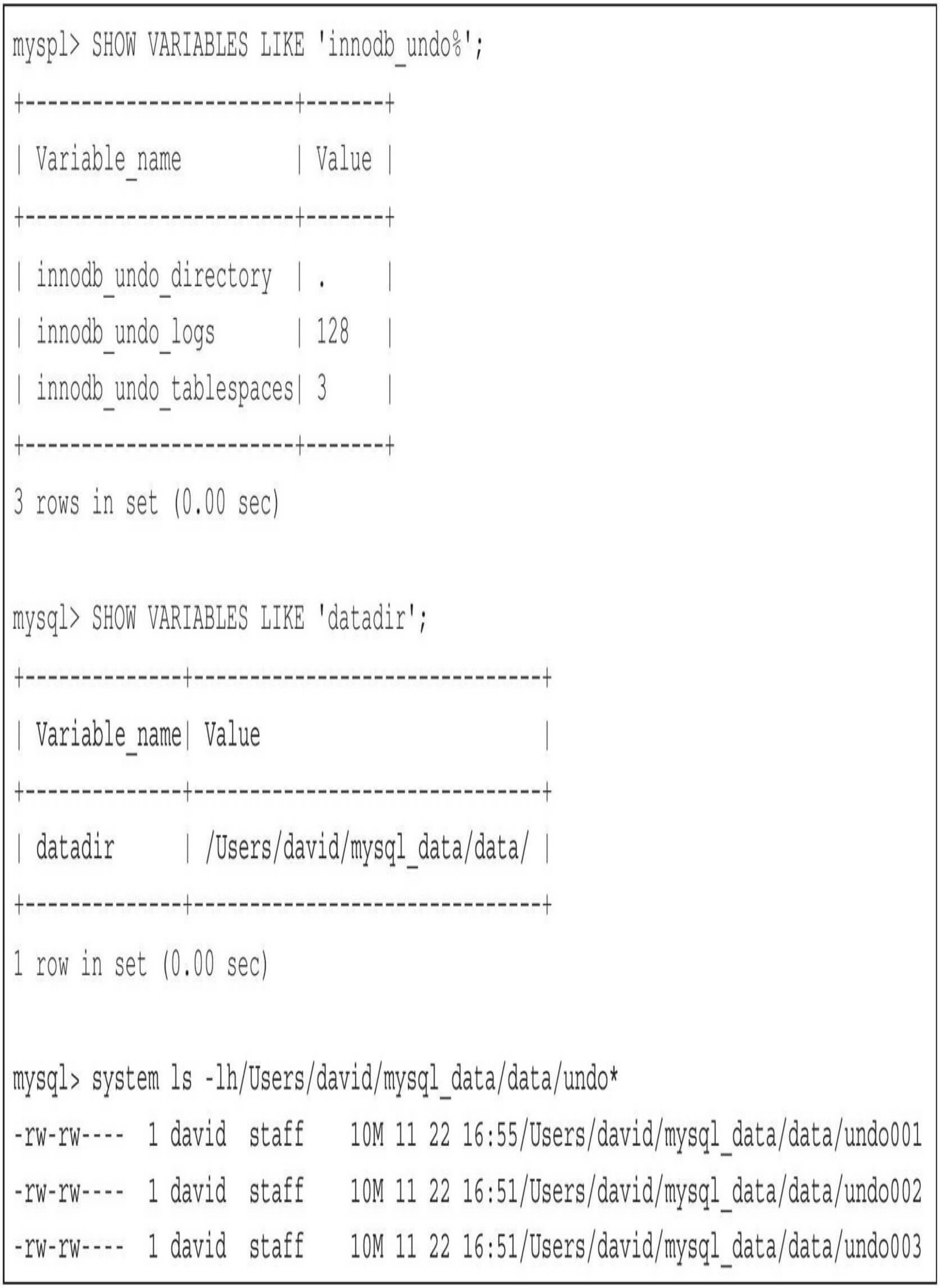
❑innodb_undo_logs

❑innodb_undo_tablespaces

□□innodb_undo_directory□□□□rollback segment□□□□□□□□□□□□rollback segment□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□“.”□□□□□InnoDB□□□□□□□□

□□innodb_undo_logs□□□□rollback segment□□□□□□□□128□□InnoDB1.2□□□□□□□□□□□□□□□□□□innodb_rollback_segments□

□□innodb_undo_tablespaces□□□□□□rollback segment□□□□□□□□rollback segment□□□□□□□□□□□□□□□□□□□□□□□□□□□innodb_undo_directory□□undo□□□□□□□□□□□□□rollback segment□□□□7-13□□□□□□□3□□□□□□rollback segment□



```
myspl> SHOW VARIABLES LIKE 'innodb_undo%';
+-----------------------+-------+
| Variable_name         | Value |
+-----------------------+-------+
| innodb_undo_directory | .     |
| innodb_undo_logs      | 128   |
| innodb_undo_tablespaces| 3    |
+-----------------------+-------+
3 rows in set (0.00 sec)

mysql> SHOW VARIABLES LIKE 'datadir';
+-------------+-------------------------------+
| Variable_name| Value                        |
+-------------+-------------------------------+
| datadir     | /Users/david/mysql_data/data/ |
+-------------+-------------------------------+
1 row in set (0.00 sec)

mysql> system ls -lh/Users/david/mysql_data/data/undo*
-rw-rw----  1 david  staff   10M 11 22 16:55/Users/david/mysql_data/data/undo001
-rw-rw----  1 david  staff   10M 11 22 16:51/Users/david/mysql_data/data/undo002
-rw-rw----  1 david  staff   10M 11 22 16:51/Users/david/mysql_data/data/undo003
```

□ 7-13 □3□□□□□□rollback segment

□□□□□□□□□□□□undo log segment□□□□□□undo log□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□

❑□undo log□□□□□□□□□□□purge□□

❑□□undo log□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□undo log□undo log□□□□□□□□□□□□□□□□□□□□□undo log□□□□□□□□□□□□□□□□□□□undo log□□□□□□□□□□□□□□□□undo log□undo log□□□□purge□□□□□□

□□□□□□□□□□□□□□□□□undo□□□□□□□□□□□□□□□□OLTP□□TPS□transaction per second□□1000□□□□□□□□□□undo□□□□□□□□□□1000*60□□□□□□□□□□□□□1GB□□□□□□purge□□□□□20□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□undo□□□□□□□□□□□□□□□□□□□□□□□undo log□□□□□□□□□□undo□□□□□□□□□□3/4□□□□□□□undo□□□□□□□□□□□undo log□□□□□undo log□□□□□□□□undo log□□□□□□□□□□□□□□undo□□□□□□□□□□□undo log□□□purge□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□SHOW ENGINE INNODB STATUS□□□□□□undo log□□□□□□

```
mysql□SHOW ENGINE INNODB STATUS\G;

***************************1.row***************************

......

------------

TRANSACTIONS

------------

Trx id counter 3000

Purge done for trx's n:o□2C03 undo n:o□0

History list length 12
```

```
LIST OF TRANSACTIONS FOR EACH SESSION:

---TRANSACTION 0,not started

MySQL thread id 1,OS thread handle 0x1500f1000,query id 4 localhost root

show engine innodb status

......
```

History list length□□□□undo log□□□□□□□12□purge□□□□□□□□□□□□undo log□□□□□□□□□□□□□□□□□□□History list length□□□□□□□0□

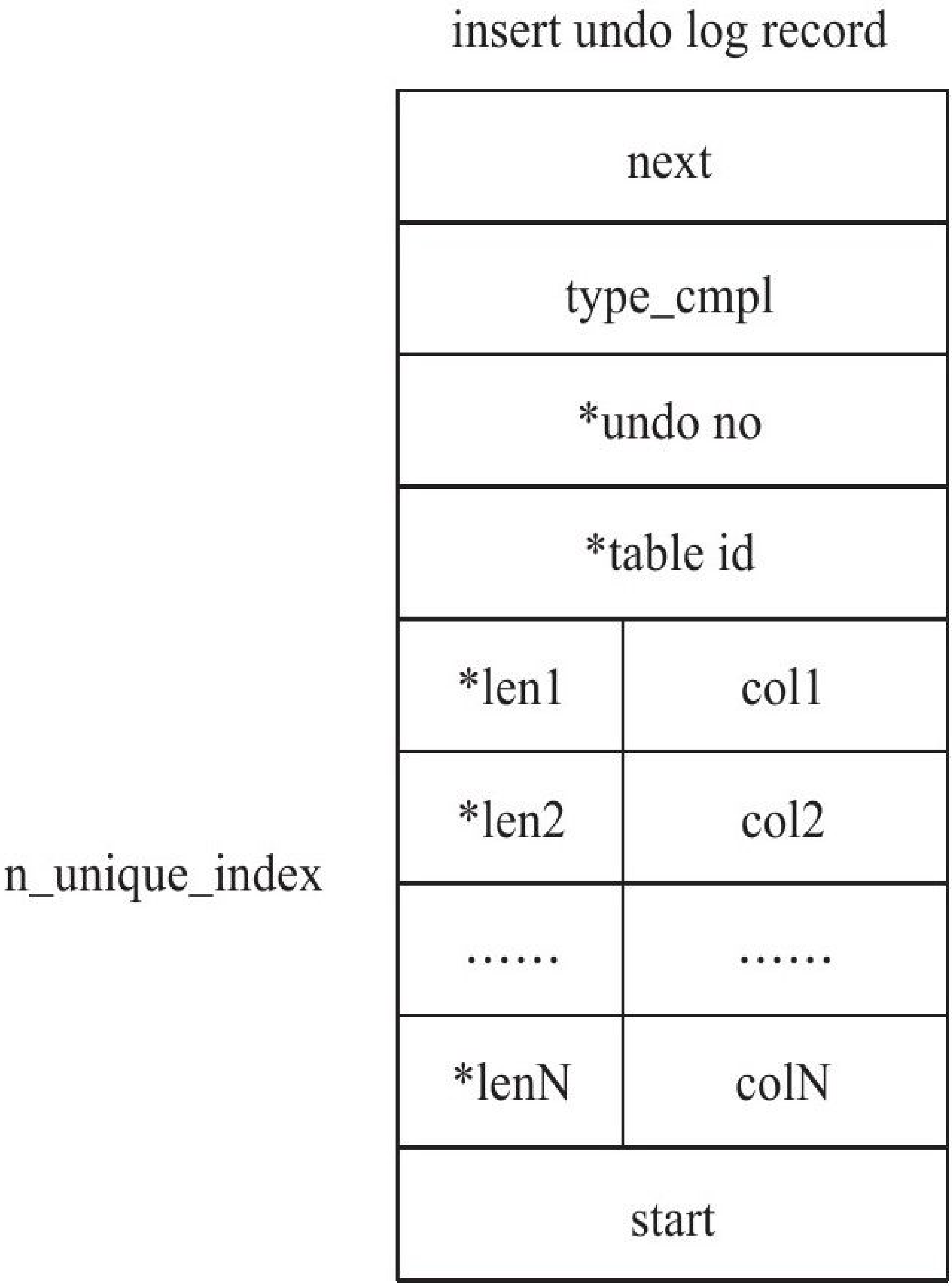
3.undo log□□

□InnoDB□□□□□□undo log□□□

❑insert undo log

❑update undo log

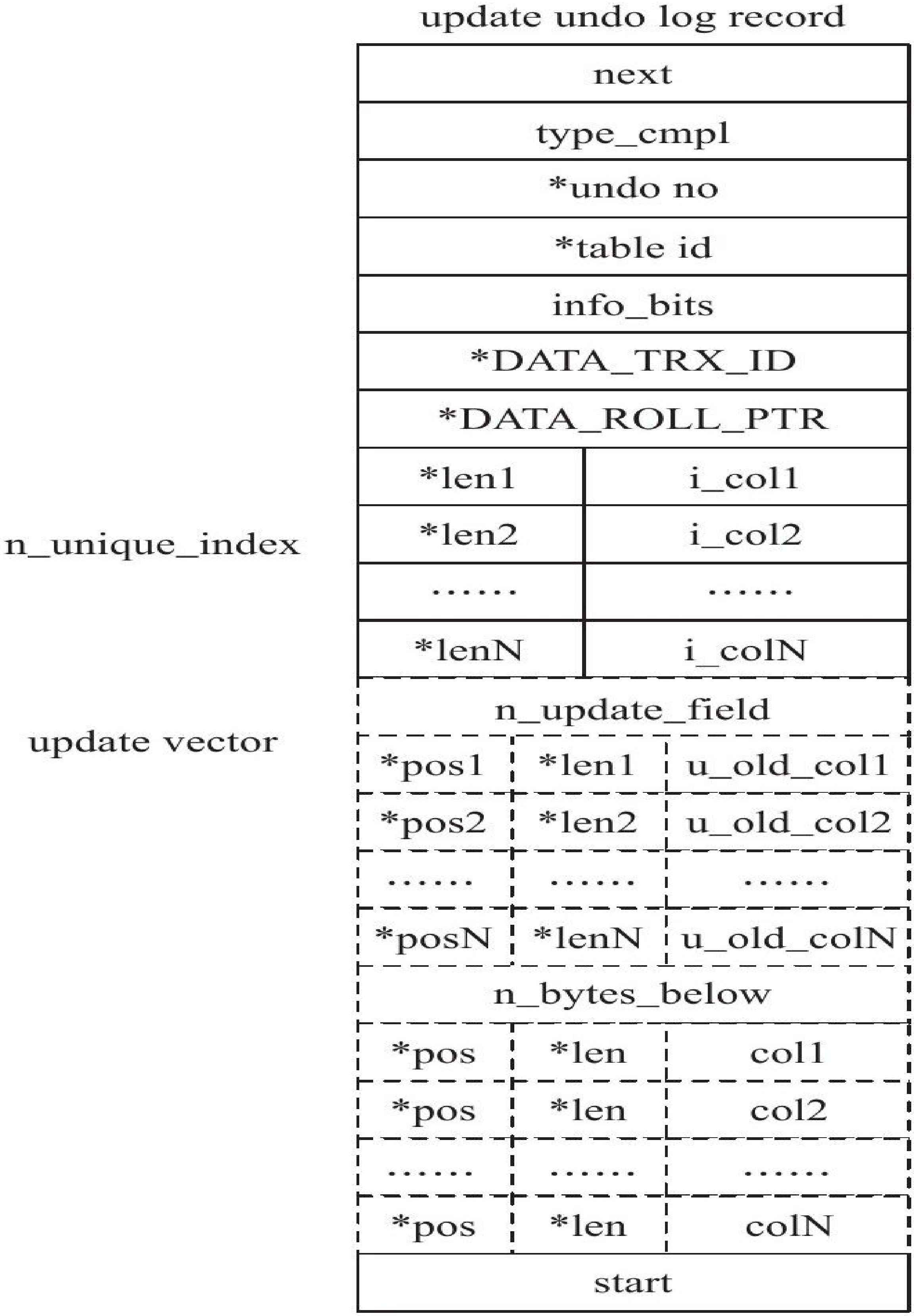
insert undo log□□□insert□□□□□□undo log□□□insert□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□undo log□□□□□□□□□□□□□□□□□□purge□□□insert undo log□□□□□7-14□□□


insert undo log record
next
type_cmpl
*undo no
*table id
*len1
col1
*len2
col2
n_unique_index
……
……
*lenN
colN
start

[illegible] 7-14 insert undo log[illegible]

[illegible]7-14[illegible]insert undo log[illegible]*[illegible]insert undo log[illegible]next[illegible]undo log[illegible]next[illegible]undo log[illegible]undo log[illegible]type_cmpl[illegible]undo[illegible]insert undo log[illegible]11[illegible]undo_no[illegible]ID[illegible]table_id[illegible]undo log[illegible]rollback[illegible]

update undo log[illegible]delete[illegible]update[illegible]undo log[illegible]undo log[illegible]MVCC[illegible]undo log[illegible]purge[illegible]update undo log[illegible]7-15[illegible]


update undo log record
next
type_cmpl
*undo no
*table id
info_bits
*DATA_TRX_ID
*DATA_ROLL_PTR
n_unique_index
*len1
i_col1
*len2
i_col2
……
……
*lenN
i_colN
update vector
n_update_field
*pos1
*len1
u_old_col1
*pos2
*len2
u_old_col2
……
……
……
*posN
*lenN
u_old_colN
n_bytes_below
*pos
*len
col1
*pos
*len
col2
……
……
……
*pos
*len
colN
start

□ 7-15 update undo log□□

update undo log□□□□□□□□insert undo log□□□□□□□□□□□□□□□□□□□□next□start□undo_no□table_id□□□□□□insert undo log□□□□□□□□type_cmpl□□□update undo log□□□□□□□□□□□□□□□□

❑12 TRX_UNDO_UPD_EXIST_REC□□non-delete-mark□□□

❑13 TRX_UNDO_UPD_DEL_REC□delete□□□□□□not delete

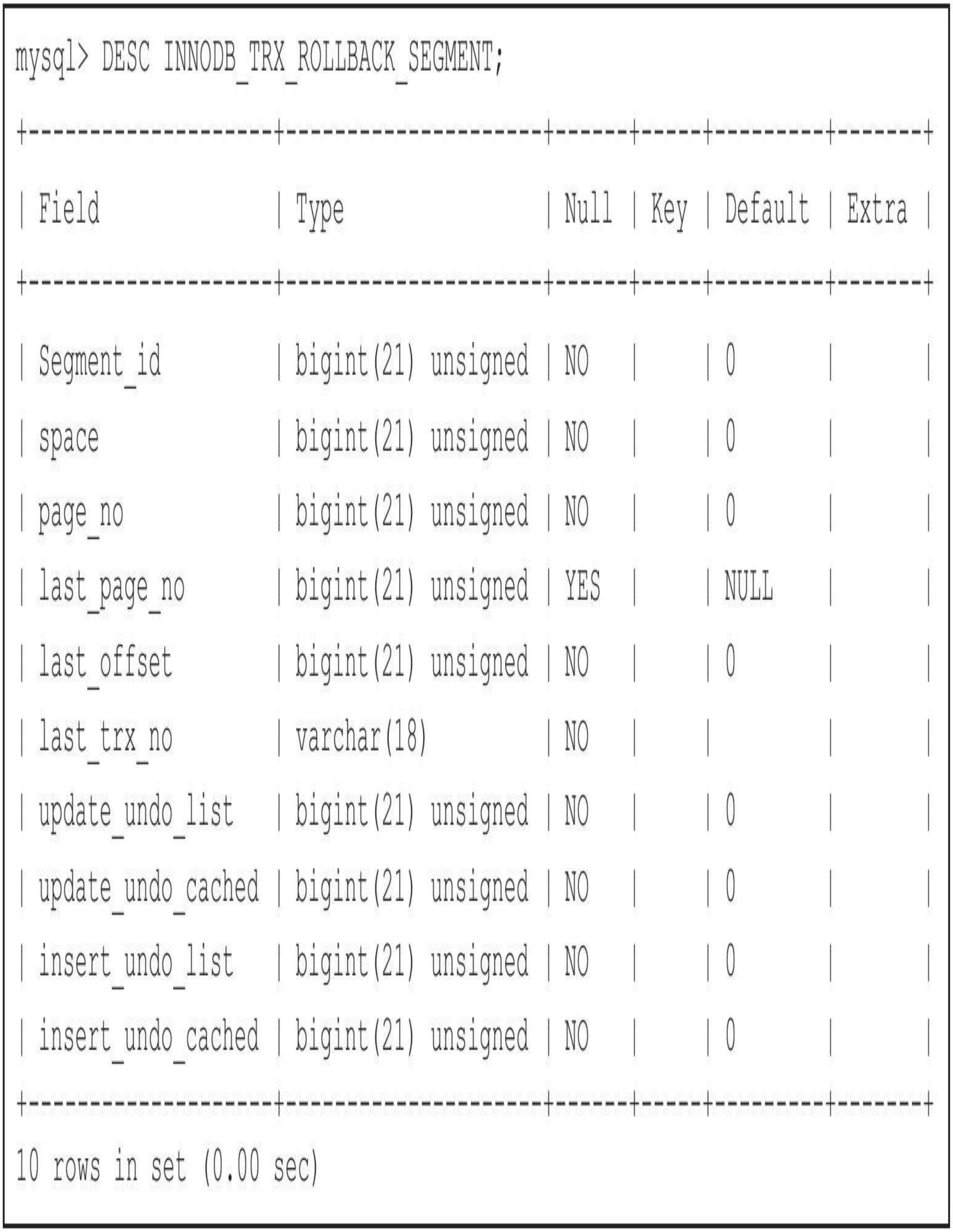
❑14 TRX_UNDO_DEL_MARK_REC□□□□□□delete

□□□□□□□update_vector□□□update_vector□□update□□□□□□□□□□□□□□□□□□□□□□□□undo log□□□□□□□undo log□□□□□□□□□□□□□□□□□□□□

4.□□undo□□

Oracle□Microsoft SQL Server□□□□□□□□□□□□□□□□□undo□□□□InnoDB□□□□□□□□□□□□□□DBA□□□□□□□□□□□□□□□InnoSQL□information_schema□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□undo□□□□

□□□□□□□□□□□INNODB_TRX_ROLLBACK_SEGMENT□□□□□□□□□□□□□□□□□rollback segment□□□□□□□7-16□□□



```
mysql> DESC INNODB_TRX_ROLLBACK_SEGMENT;
+--------------------+---------------------+------+-----+---------+-------+
| Field              | Type                | Null | Key | Default | Extra |
+--------------------+---------------------+------+-----+---------+-------+
| Segment_id         | bigint(21) unsigned | NO   |     | 0       |       |
| space              | bigint(21) unsigned | NO   |     | 0       |       |
| page_no            | bigint(21) unsigned | NO   |     | 0       |       |
| last_page_no       | bigint(21) unsigned | YES  |     | NULL    |       |
| last_offset        | bigint(21) unsigned | NO   |     | 0       |       |
| last_trx_no        | varchar(18)         | NO   |     |         |       |
| update_undo_list   | bigint(21) unsigned | NO   |     | 0       |       |
| update_undo_cached | bigint(21) unsigned | NO   |     | 0       |       |
| insert_undo_list   | bigint(21) unsigned | NO   |     | 0       |       |
| insert_undo_cached | bigint(21) unsigned | NO   |     | 0       |       |
+--------------------+---------------------+------+-----+---------+-------+
10 rows in set (0.00 sec)
```

□ 7-16 INNODB_TRX_ROLLBACK_SEGMENT□□□

□□□□□□□□□□□□□□□rollback segment□□□□□

```
mysql□SELECT segment_id,space,page_no

-□FROM INNODB_TRX_ROLLBACK_SEGMENT;

+------------+-------+---------+

|segment_id|space|page_no|

+------------+-------+---------+

|0|0|6|

|1|0|45|

|2|0|46|

......

128 rows in set(0.00 sec)
```

□□□□□□□□□INNODB_TRX_UNDO□□□□□□□□□□undo log□□□DBA□□□□□□□□□□□□□□□□undo□□□□□□□□□□□INNODB_TRX_UNDO□□□□□□□□□□□□□□□t□

```
CREATE TABLE t(

a INT,

b VARCHAR(32),

PRIMARY KEY(a),

KEY(b)

)ENGINE=InnoDB;
```

□□□□□□□□□□□□□□INNODB_TRX_UNDO□□□□□□undo log□□□□

```
mysql□TBEGIN;

Query OK,0 rows affected(0.00 sec)

mysql□INSERT INTO t SELECT 1,'1';

Query OK,1 row affected(0.00 sec)
```

```
Records:1 Duplicates:0 Warnings:0

mysql□SELECT*FROM information_schema.INNODB_TRX_UNDO\G;

***************************1.row***************************

trx_id:3001

rseg_id:2

undo_rec_no:0

undo_rec_type:TRX_UNDO_INSERT_REC

size:12

space:0

page_no:334

offset:272

1 row in set(0.00 sec)
```

□□□□□□□□□□□□□□□ID□3001□rollback segment□ID□2□□□□□□□□□□□□□□□□□undo_rec_no□0□□□□□□□□□□□□□□TRX_UNDO_INSERT_REC□□□□insert undo log□size□□undo log□□□□□□12□□□□□□space□page_no□offset□□undo log□□□□□□□□□□□ibdata1□□□□□□334□272□□□□□12□□□□□□□□□□□□

```
01 1c 0b 00 16 04 80 00 00 01 01 10
```

□□□□undo log□□□□□□□□□□□□□□undo log□□□□□□□□□□□□□□

```
01 1c#□□□undo log□□□272+12=0x011c

0b#undo log□□□□TRX_UNDO_INSERT_REC□11

00#undo log□□□□□□□undo_rec_no

16#□□ID

04#□□□□□

80 00 00 01#□□□□□

01 10#undo log□□□□□□□272=0x0110
```

□□□□□□□□undo log□□□rollback segment□ID□2□□□□□□□□□□□□□INNODB_TRX_ROLLBACK_SEGMENT□□□□□rollback segment□□□□□□

```
mysql□SELECT segment_id,insert_undo_list,insert_undo_cached

-□FROM information_schema.INNODB_TRX_ROLLBACK_SEGMENT

-□WHERE segment_id=2\G;

***************************1.row***************************

segment_id:2

insert_undo_list:1

insert_undo_cached:0

1 row in set(0.00 sec)
```

□□□□insert_undo_list□1□□□□□□□□□COMMIT□□□□□□□□□□□□□

```
mysql□COMMIT;

Query OK,0 rows affected(0.00 sec)

mysql□SELECT segment_id,insert_undo_list,insert_undo_cached

-□FROM information_schema.INNODB_TRX_ROLLBACK_SEGMENT

-□WHERE segment_id=2\G;

***************************1.row***************************

segment_id:2

insert_undo_list:0

insert_undo_cached:1

1 row in set(0.00 sec)
```

□□□□□insert_undo_list□□0□□insert_undo_cached□□□1□□□□□□□□□□undo□□□□□□□□□□□□□□rollback segment□□undo□□□□□□□□□□□□

□□□□□□delete□□□□□undo log□□□□□□□□

```
mysql□BEGIN;

Query OK,0 rows affected(0.00 sec)

mysql□DELETE FROM t WHERE a=1;

Query OK,1 row affected(0.00 sec)

Records:1 Duplicates:0 Warnings:0

mysql□SELECT*FROM information_schema.INNODB_TRX_UNDO\G;

***************************1.row***************************

trx_id:3201

rseg_id:2

undo_rec_no:0

undo_rec_type:TRX_UNDO_DEL_MARK_REC

size:37

space:0

page_no:326

offset:620

1 row in set(0.00 sec)
```

□□□□□□□□□□□□326□□□□□620□□□□□□□□□□□

```
0518260 00 00 00 00 00 00 00 00 00 00 00 00 02 91 0e 00

0518270 16 00 00 00 00 30 01 e0 82 00 00 01 4e 01 10 04

0518280 80 00 00 01 00 0b 00 04 80 00 00 01 03 01 31 02

0518290 6c 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
```

□□□□□□□

```
02 91#□□□undo log□□□□□□□

0e#undo log□□□TRX_UNDO_DEL_MARK_REC□14

00#undo no
```

```
16#table id

00#info bits

00 00 00 30 01 e0#rec□□id

82 00 00 01 4e 01 10#rec□□□□

04#□□□□

80 00 00 01#□□□

00 0b#□□□□□□□

00#□□□□

04#□□□□

80 00 00 01#□□□

03#□□□□□□00□02□□□□

01#□□□□

31#□b□□□□□□'1'□□□□□

02 6c#□□□□□□□□
```

□□rollback segment□□□□□□□□

```
mysql□SELECT segment_id,update_undo_list,update_undo_cached

-□FROM information_schema.INNODB_TRX_ROLLBACK_SEGMENT

-□WHERE segment_id=2\G;

***************************1.row***************************

segment_id:2

update_undo_list:1

update_undo_cached:0

1 row in set(0.00 sec)
```

□□□□□□□□□□□undo□□□□cache□□□□□□□□□

```
mysql□COMMIT;

Query OK,0 rows affected(0.00 sec)
```

```
mysql□SELECT segment_id,update_undo_list,update_undo_cached

-□FROM information_schema.INNODB_TRX_ROLLBACK_SEGMENT

-□WHERE segment_id=2\G;

***************************1.row***************************

segment_id:2

update_undo_list:0

update_undo_cached:1

1 row in set(0.00 sec)
```

[illegible]delete[illegible]delete flag[illegible]1[illegible]purge[illegible]

[illegible]update[illegible]undo log[illegible]1[illegible]'1'[illegible]update[illegible]INNODB_TRX_UNDO[illegible]undo log[illegible]

```
mysql□INSERT INTO t SELECT 1,'1';

mysql□BEGIN;

Query OK,0 rows affected(0.00 sec)

mysql□UPDATE t SET b='2'WHERE a=1;

Query OK,1 row affected(0.00 sec)

Rows matched:1 Changed:1 Warnings:0

mysql□SELECT*FROM information_schema.INNODB_TRX_UNDO\G;

***************************1.row***************************

trx_id:3205

rseg_id:5

undo_rec_no:0

undo_rec_type:TRX_UNDO_UPD_EXIST_REC

size:41

space:0

page_no:318

offset:724
```

```
1 row in set(0.00 sec)
```

□□□□□□□□□□□□318□□□□□724□□□□□□□□□□□

```
04f82d0 00 00 00 00 02 fd 0c 00 16 00 00 00 00 32 04 e0

04f82e0 84 00 00 01 48 01 10 04 80 00 00 01 01 03 01 31

04f82f0 00 0b 00 04 80 00 00 01 03 01 31 02 d4 00 00 00
```

□□□□□□

```
02 fd#□□□undo log□□□□□

0c#undo log□□□TRX_UNDO_UPD_DEL_REC□13

00#undo no

16#table id

00#info bits

00 00 00 32 04 e0#rec trx id

84 00 00 01 48 01 10#rec□□□□

04#□□□□

80 00 00 01#□□□

01#update vector□□□

03#update vector□b□□□

01#update vector□□□□

31#update vector□□□□□□□'1'

00 0b#□□□□□□□□□□

00#□□□□

04#□□□□

80 00 00 01#□□□

03#□□□□

31#□□□

02 d4#undo log□□□□□□□□
```

□□□□□□□□□□□□□□□□□□□□□□□□□□□□undo log□□□□□□

```
mysql□ROLLBACK;

Query OK,1 row affected(0.00 sec)

mysql□UPDATE t SET a=2 WHERE a=1;

Rows matched:1 Changed:1 Warnings:0

mysql□SELECT*FROM information_schema.INNODB_TRX_UNDO

-□ORDER BY undo_rec_no\G;

***************************1.row***************************

trx_id:320F

rseg_id:11

undo_rec_no:0

undo_rec_type:TRX_UNDO_DEL_MARK_REC

size:37

space:0

page_no:324

offset:492

***************************2.row***************************

trx_id:320F

rseg_id:11

undo_rec_no:1

undo_rec_type:TRX_UNDO_INSERT_REC

size:12

space:0

page_no:336

offset:272

2 rows in set(0.00 sec)
```

[illegible]update[illegible]TRX_UNDO_DEL_MARK_REC[illegible]undo log[illegible]TRX_UNDO_INSERT_REC[illegible]undo log[illegible]undo_rec_no[illegible]undo log[illegible]

[illegible]InnoSQL[illegible]undo[illegible]DBA[illegible]undo[illegible]

7.2.3 purge

delete[illegible]update[illegible]t[illegible]SQL[illegible]

```
DELETE FROM t WHERE a=1;
```

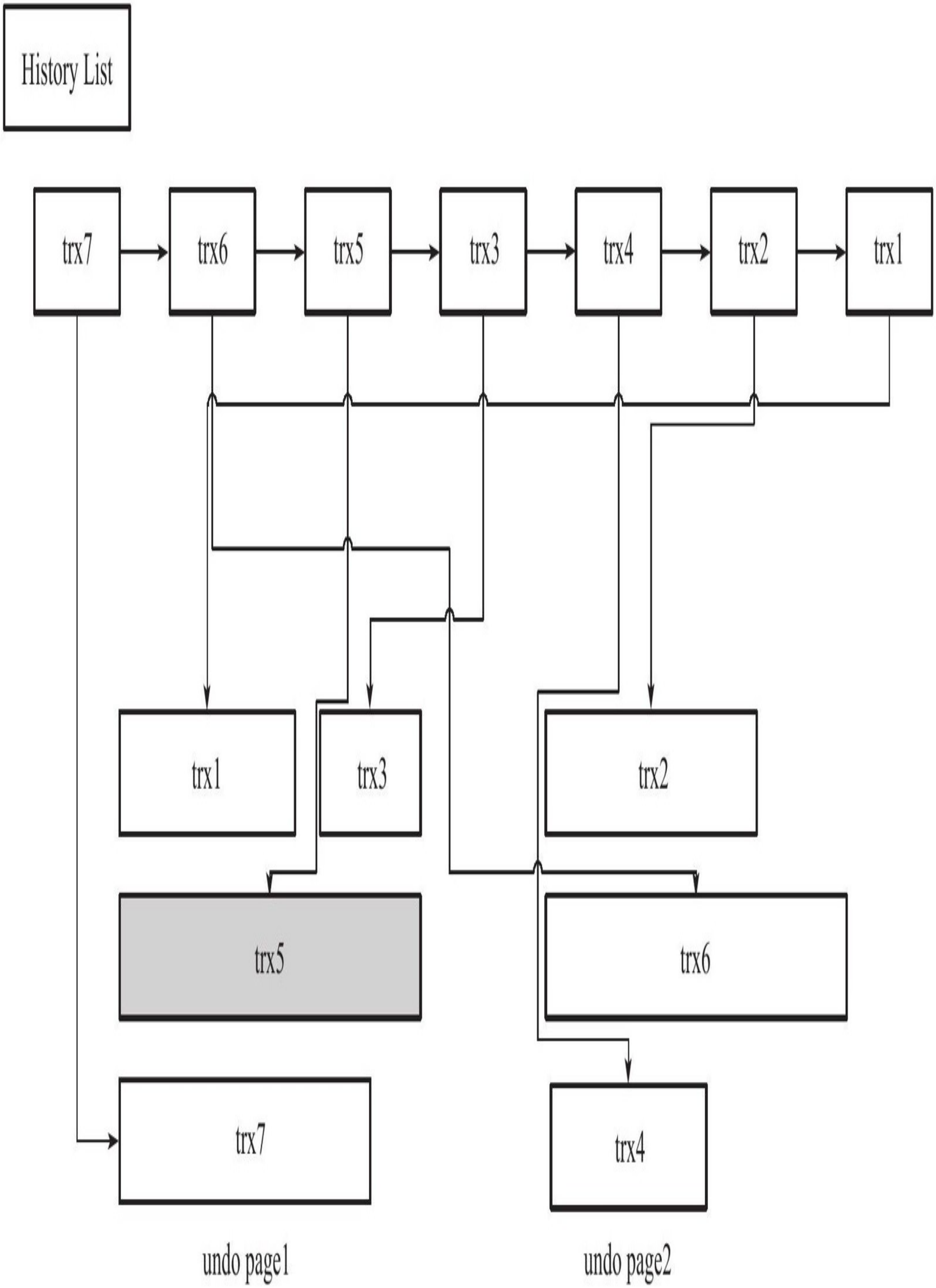
[illegible]t[illegible]a[illegible]b[illegible]delete[illegible]undo log[illegible]1[illegible]delete flag[illegible]1[illegible]B+[illegible]a[illegible]1[illegible]b[illegible]1[illegible]undo log[illegible]“[illegible]”[illegible]purge[illegible]

purge[illegible]delete[illegible]update[illegible]InnoDB[illegible]MVCC[illegible]InnoDB[illegible]purge[illegible]delete[illegible]purge[illegible]delete[illegible]update[illegible]“[illegible]”[illegible]delete[illegible]

[illegible]InnoDB[illegible]undo log[illegible]undo log[illegible]undo log[illegible]InnoDB[illegible]history[illegible]undo log[illegible]

[illegible]7-17[illegible]history list[illegible]undo log[illegible]InnoDB[illegible]undo page[illegible]undo log[illegible]undo page[illegible]undo log[illegible]trx5[illegible]undo log[illegible]

[illegible]purge[illegible]InnoDB[illegible]history list[illegible]trx1[illegible]InnoDB[illegible]trx1[illegible]undo log[illegible]trx3[illegible]trx5[illegible]trx5[illegible]history list[illegible]trx2[illegible]trx2[illegible]trx6[illegible]trx4[illegible]undo page2[illegible]undo page[illegible]


History List
trx7
trx6
trx5
trx3
trx4
trx2
trx1
trx1
trx3
trx2
trx5
trx6
trx7
trx4
undo page1
undo page2

□ 7-17 undo log□history□□□□□

InnoDB□□□□□□□□history list□□undo log□□□□□undo page□□undo log□□□□□□□□□□□□□□□□□□□□□□□□purge□□□□

□□□□□□innodb_purge_batch_size□□□□□□purge□□□□□□□undo page□□□□InnoDB1.2□□□□□□□□□□□20□□□1.2□□□□□□□□□□□□□300□□□□□□□□□□□□□□□□□□□□undo page□□□□□□□□□□□□undo page□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□purge□□□□□undo page□□□□□CPU□□□IO□□□□□□undo log□□□□□□□□□□□□□□□□□□□□□□□□□□DBA□□□□□□□□□□□□□□□□□□□□□□□□□□□MySQL□□□□□□□□□□□□□□□□□□□□□□

□InnoDB□□□□□□□□□□□□□□□□□□□□purge□□□□□history list□□□□□□□□□□□□□□□□□innodb_max_purge_lag□□□□history list□□□□□□□□□□□□□□□□“□□”DML□□□□□□□□□□□0□□□□□history list□□□□□□□□□0□□□□□□DML□□□□□□□□□□□□

```
delay=□□length(history_list)-innodb_max_purge_lag□*10)-5
```

delay□□□□□□□□□□□□□□□□□□□delay□□□□□□□□□□□DML□□□□□□□□update□□□□□□5□□□□□□□□□□□□□□□delay□□□□□□□□□5*delay□□delay□□□□□□□□purge□□□□□□□□□□□□□

InnoDB1.2□□□□□□□□□□□□□innodb_max_purge_lag_delay□□□□□□delay□□□□□□□□□□□□□□□□□□delay□□□□□□□□□delay□□□innodb_max_purge_lag_delay□□□□□purge□□□□□□□□SQL□□□□□□□□□□□

7.2.4 group commit

[illegible]fsync[illegible]fsync[illegible]fsync[illegible]group commit[illegible]fsync[illegible]InnoDB[illegible]

1[illegible]

2[illegible]fsync[illegible]

[illegible]2[illegible]1[illegible]1[illegible]2[illegible]fsync[illegible]group commit[illegible]

[illegible]InnoDB1.2[illegible]InnoDB[illegible]group commit[illegible]replication[illegible]

[illegible]

1[illegible]InnoDB[illegible]prepare[illegible]

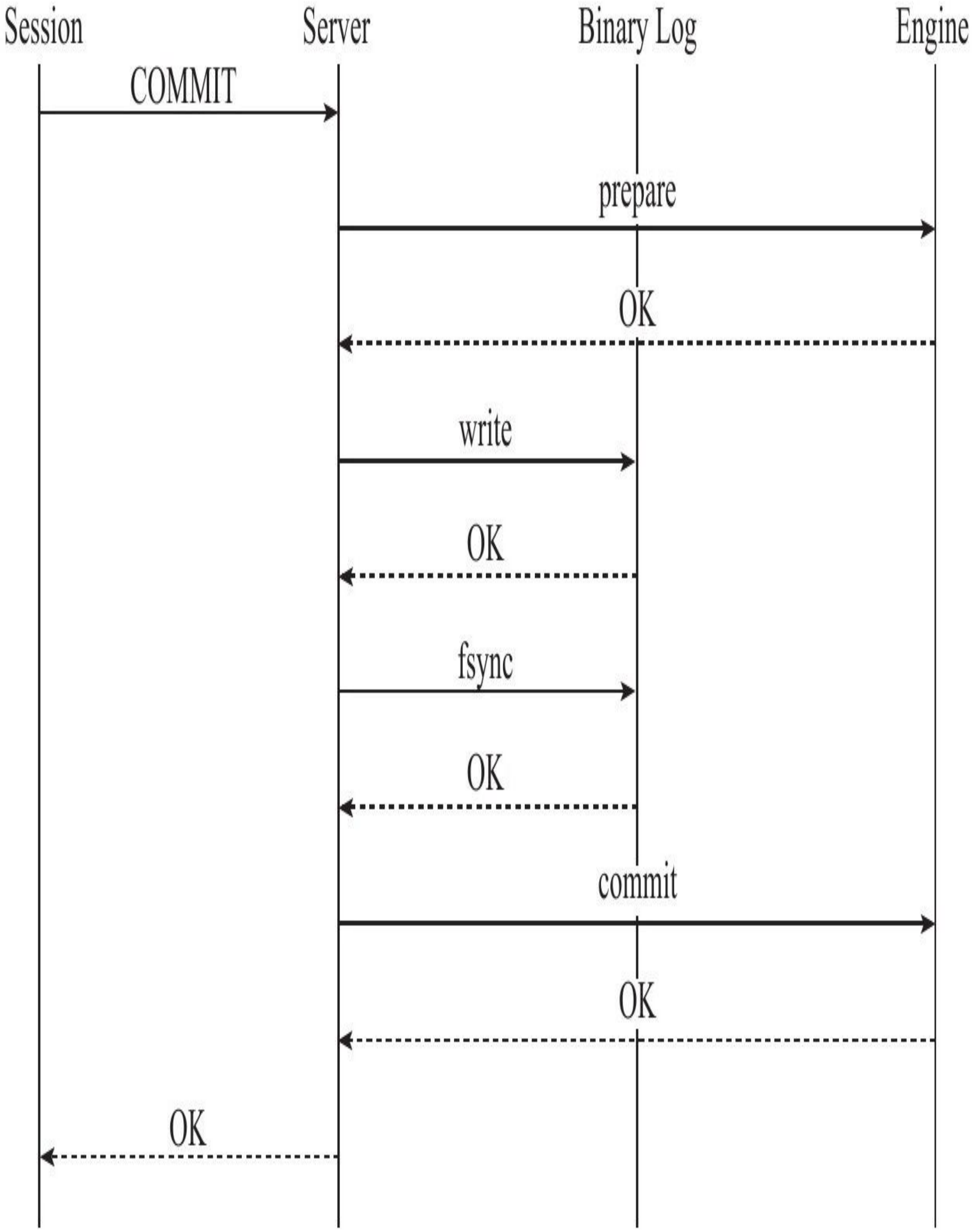
2[illegible]MySQL[illegible]

3[illegible]InnoDB[illegible]

a[illegible]

b[illegible]fsync[illegible]

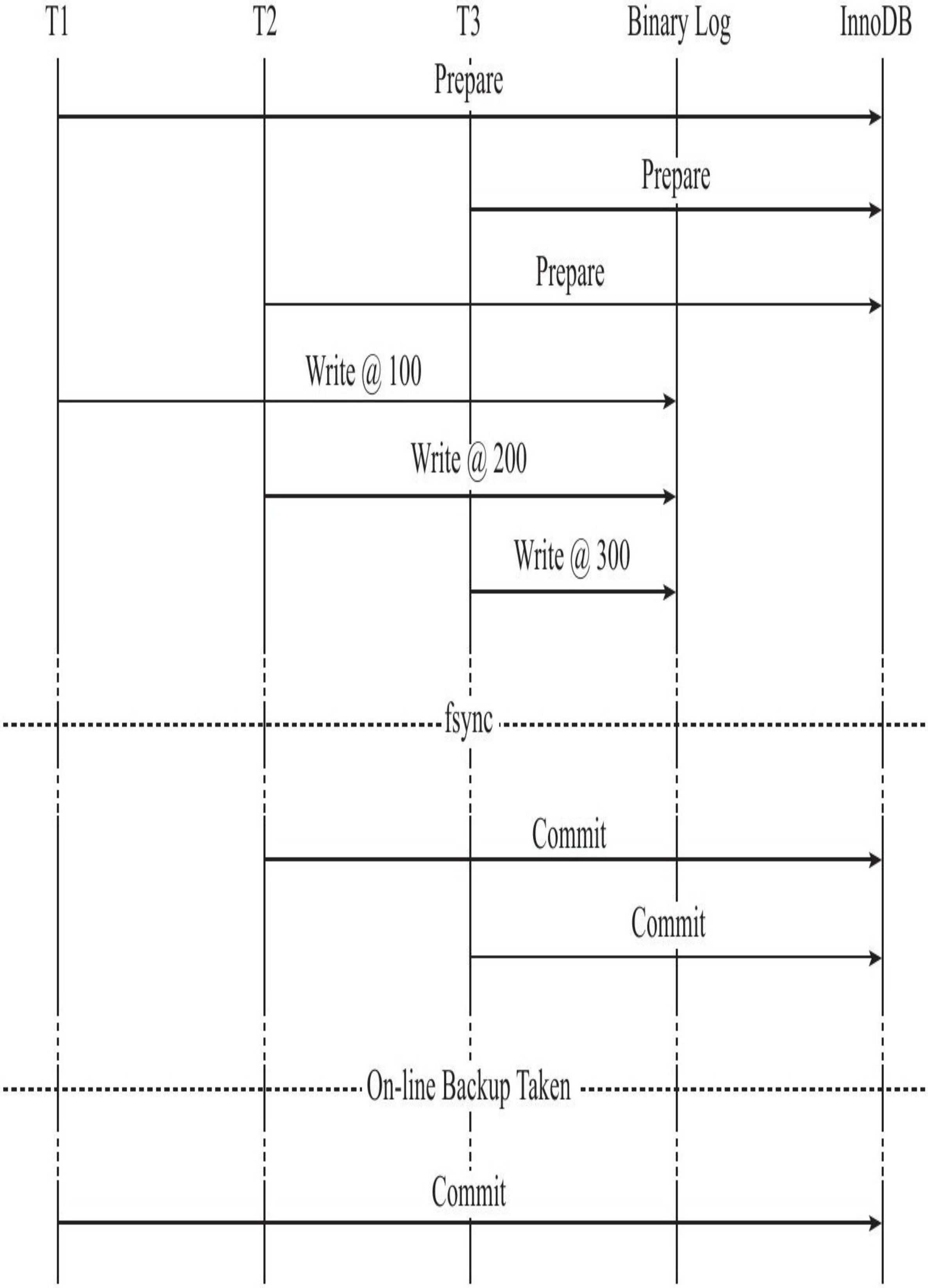
[illegible]2[illegible]3[illegible]fsync[illegible]2[illegible]fsync[illegible]sync_binlog[illegible]3[illegible]fsync[illegible]innodb_flush_log_at_trx_commit[illegible]7-18[illegible]


Session
Server
Binary Log
Engine
COMMIT
prepare
OK
write
OK
fsync
OK
commit
OK
OK

□ 7-18 □□□□□□□□InnoDB□□□□□□□□□

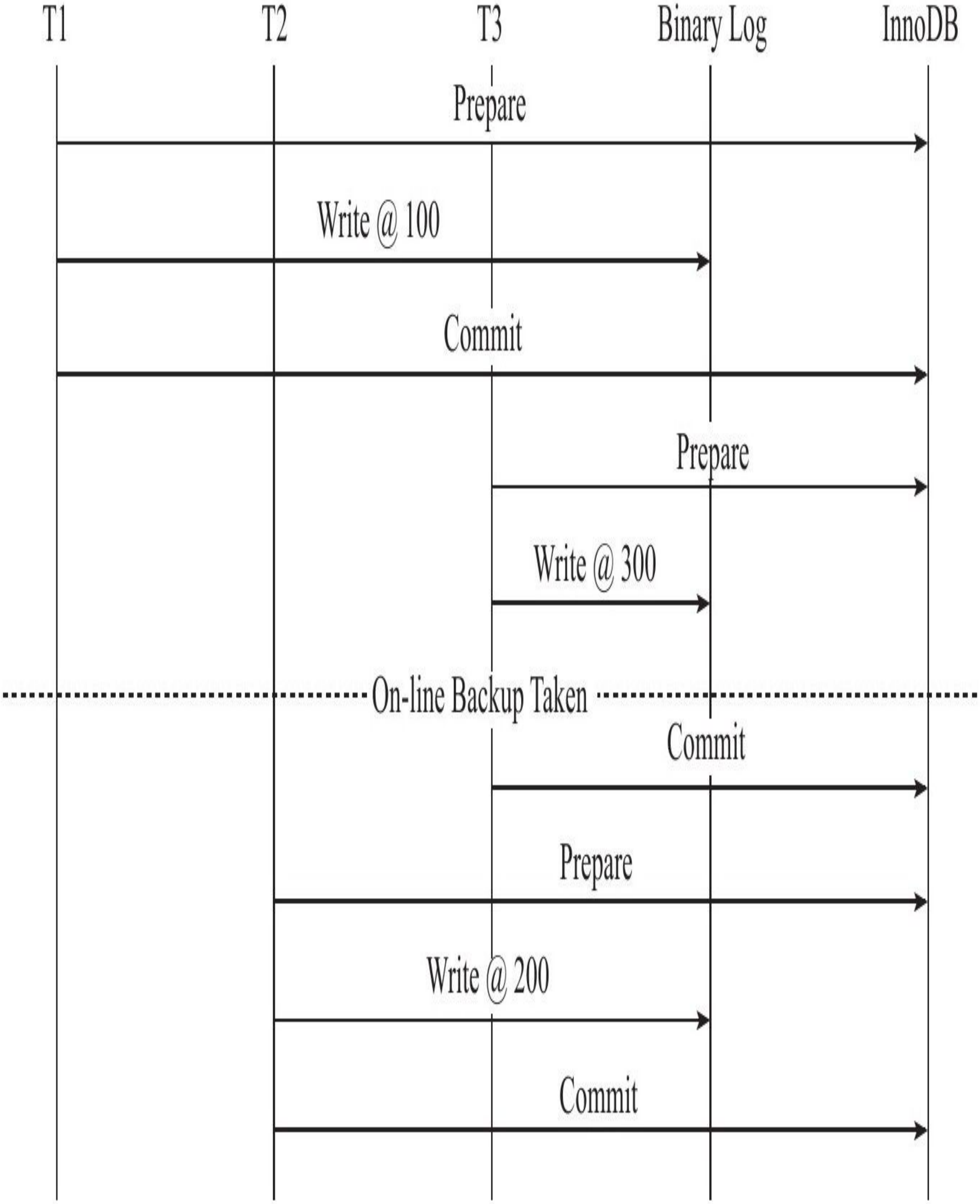
□□□□MySQL□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□MySQL□□□□□□□□prepare_commit_mutex□□□□□□□□□□□□□□□□□3□□□□□a□□□□□□□□□□□□□□b□□□□□□□□□□group commit□□□

□□□□□□□□□□MySQL□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□xtrabackup□□ibbackup□□□□□□□□□□replication□□□7-19□□□


T1
T2
T3
Binary Log
InnoDB
Prepare
Prepare
Prepare
Write @ 100
Write @ 200
Write @ 300
fsync
Commit
Commit
On-line Backup Taken
Commit

□ 7-19 InnoDB□□□□□□□□□□□□□MySQL□□□□□□□□□□□□□

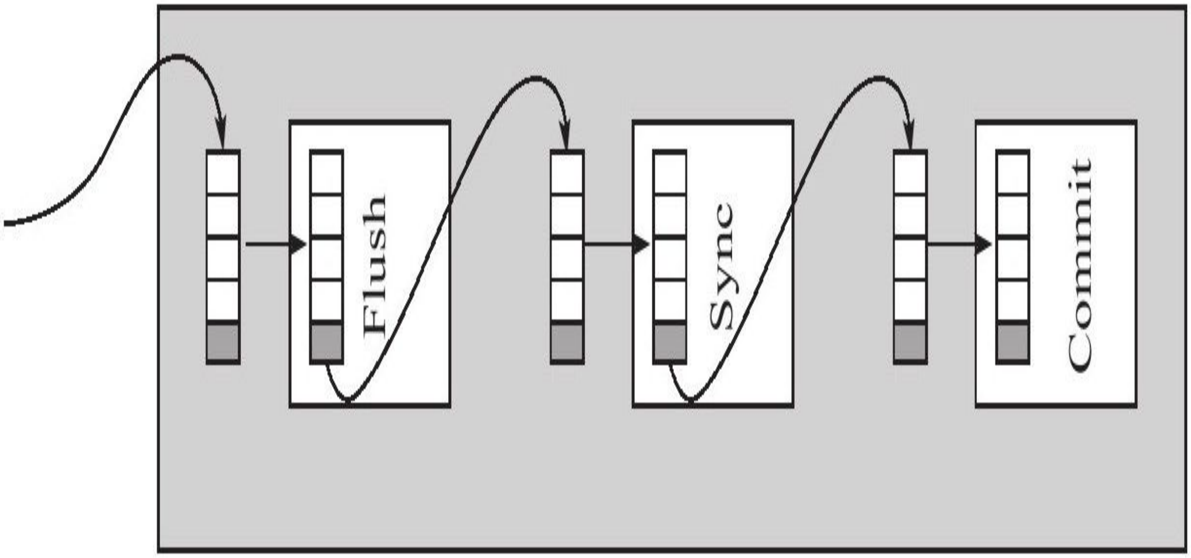
□□□□□□□□□□□□□□□□□□□□□□□□□□□replication□□□T1□□□□□□□□□□□□InnoDB□□□□□□□□□□T3□□□□□□□□□□□□□□□□□□□□□□□□□□prepare_commit_mutex□□□□□□□□□□□□□□□□□□group commit□□□□□□□7-20□□□


T1
T2
T3
Binary Log
InnoDB
Prepare
Write @ 100
Commit
Prepare
Write @ 300
On-line Backup Taken
Commit
Prepare
Write @ 200
Commit

□ 7-20 □□□prepare_commit_mutex□□InnoDB□□□□□□□□□□
MySQL□□□□□□□□□□□□□□□□□

□□□□□□□2010□□MySQL□□□□□□□□□Facebook MySQL□□□□Percona□□□□□□□□□□□□□□MariaDB□□□□□□□□Kristian Nielsen□□□□□□“□□”□□□□□□□□□□□□□□MySQL□□□□□□□□□□□□□□group commit□□InnoDB□□□□□□□group commit□□□□□□□□□□□□prepare_commit_mutex□□□□□□□□□□□□□□□□MySQL 5.6□□□□□□□□□□□□□□□□Binary Log Group Commit□BLGC□□

MySQL 5.6 BLGC□□□□□□□□□□□□□□□□□□□□□□□□□□7-21□□□



□ 7-21 MySQL 5.6 BLGC□□□□□

□MySQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□leader□□□□□□□follower□leader□□□follower□□□□BLGC□□□□□□□□□□□□

❑Flush□□□□□□□□□□□□□□□□□□□□

❑Sync□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□fsync□□□□□□□□□□□□□□□□□□BLGC□

❏Commit□□□leader□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□group commit□□□□□□□□□□□prepare_commit_mutex□□group commit□□□□□□

□□□□□□□□□Commit□□□□□□□□□□□□□Flush□□□□□□group commit□□□□□□□group commit□□group commit□□□□□□□□□□□□□□□□□□□□

□□binlog_max_flush_queue_time□□□□Flush□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Sync□□□□□□□□□□□□□□□□□□□□□□□group commit□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□0□□□□□□□□□0□□□□□□MySQL□□□□□□□□□□□□□□□100□□□□□□□□□□□□□□□□□□□□□□□□

7.3 □□□□□□

□MySQL□□□□□□□□□□□□□□□□□□□auto commit□□□□□□SQL□□□□□□□□□COMMIT□□□□□□□□□□□□□□□□□□□□BEGIN□START TRANSACTION□□□□□□□SET AUTOCOMMIT=0□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□DBA□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□MySQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

❑START TRANSACTION|BEGIN□□□□□□□□□□□

❑COMMIT□□□□□□□□□□□□□□□□□□□COMMIT□□□□□□□□□□□□COMMIT WORK□□□□□□□□□□□□□COMMIT□□□□□□□□□□□□□□□□□□□□□□□□□□□

❑ROLLBACK□□□□□□□□□□□□□□□□□□□ROLLBACK□□□□□□□□□□ROLLBACK WORK□□

❑SAVEPOINT identifier:SAVEPOINT□□□□□□□□□□□□□□□□□□□□□□□□SAVEPOINT□

❑RELEASE SAVEPOINT identifier□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

❑ROLLBACK TO[SAVEPOINT]identifier□□□□□□SAVEPOINT□□□UPDATE□□□□□□□□SAVEPOINT□□□□□□□DELETE□□□□□□□DELETE□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ROLLBACK TO SAVEPOINT□□□□□□□□□□□□□SAVEPOINT□□□DELETE□□□□□□□□□UPDATE□□□□□□□□□□□□

❑SET TRANSACTION□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□READ UNCOMMITTED□READ COMMITTED□REPEATABLE READ□SERIALIZABLE□

START TRANSACTION□BEGIN□□□□□□MySQL□□□□□□□□□□□□□□□□□□□□□□□MySQL□□□□□□□□□□□BEGIN□□□BEGIN…END□□□□□□□□□□□□□START TRANSACTION□□□□□□□□□□

COMMIT□COMMIT WORK□□□□□□□□□□□□□□□□□□□□□□□□COMMIT WORK□□□□□□□□□□□□□CHAIN□□RELEASE□□□□□CHAIN□□□□□□□□□□□□□□□

□□□□□□□□completion_type□□□□□□□□□□□□0□□□□□□□□□□□□□□□□COMMIT□COMMIT WORK□□□□□□□□□□completion_type□□□1□□COMMIT WORK□□□COMMIT AND CHAIN□□□□□□□□□□□□□□□□□□□□□□□

```
mysql□CREATE TABLE t(a INT,PRIMARY KEY(a))ENGINE=INNODB;

Query OK,0 rows affected(0.00 sec)

mysql□SELECT@@autocommit\G;

***************************1.row***************************

@@autocommit:1

1 row in set(0.00 sec)

mysql□SET@@completion_type=1;

Query OK,0 rows affected(0.00 sec)

mysql□BEGIN;

Query OK,0 rows affected(0.00 sec)

mysql□INSERT INTO t SELECT 1;

Query OK,1 row affected(0.00 sec)

Records:1 Duplicates:0 Warnings:0

mysql□COMMIT WORK;

Query OK,0 rows affected(0.01 sec)

mysql□INSERT INTO t SELECT 2;

Query OK,1 row affected(0.00 sec)

Records:1 Duplicates:0 Warnings:0

mysql□INSERT INTO t SELECT 2;
```

```
ERROR 1062(23000):Duplicate entry'2'for key'PRIMARY'

mysql□ROLLBACK;

Query OK,0 rows affected(0.00 sec)

#□□□□□□□□1□□□□□□□□2□□□□

mysql□SELECT*FROM t\G;

***************************1.row***************************

a:1

1 row in set(0.00 sec)
```

□□□□□□□□□□completion_type□1□□□□□□COMMIT WORK□□□1□□□□□□□□□□2□□□□□□□BEGIN□□□START TRANSACTION□□□□□□□□□□□□□□□□□□□□□□□□2□□□□□□□□□□□ROLLBACK□□□□□□□□□□1□□□□□2□□□□□□□□□completion_type□1□□COMMIT WORK□□□□□□□□□□□□□□INSERT INTO t SELECT 2□□□□□□□□□□□□□□□□□2□□□□□□□□□□□t□□

□□completion_type□2□□COMMIT WORK□□□COMMIT AND RELEASE□□□□□□□□□□□□□□□□□□□□□□

```
mysql□SET@@completion_type=2;

Query OK,0 rows affected(0.00 sec)

mysql□BEGIN;

Query OK,0 rows affected(0.00 sec)

mysql□INSERT INTO t SELECT 3;

Query OK,1 row affected(0.00 sec)

Records:1 Duplicates:0 Warnings:0

mysql□COMMIT WORK;

Query OK,0 rows affected(0.01 sec)

mysql□SELECT@@version\G;

ERROR 2006(HY000):MySQL server has gone away

No connection.Trying to reconnect...
```

```
Connection id:54

Current database:test

***************************1.row***************************

@@version:5.1.45-log

1 row in set(0.00 sec)
```

[illegible]completion_type[illegible]2[illegible]COMMIT WORK[illegible]SELECT@@version[illegible]ERROR 2006[illegible]HY000[illegible]MySQL server has gone away[illegible]COMMIT WORK[illegible]

ROLLBACK[illegible]ROLLBACK WORK[illegible]COMMIT[illegible]COMMIT WORK[illegible]

SAVEPOINT[illegible]ROLLBACK TO SAVEPOINT[illegible]

```
mysql□BEGIN;

Query OK,0 rows affected(0.00 sec)

mysql□ROLLBACK TO SAVEPOINT t1;

ERROR 1305(42000):SAVEPOINT t1 does not exist
```

InnoDB[illegible]

```
mysql□CREATE TABLE t(a INT,PRIMARY KEY(a))ENGINE=INNODB;

Query OK,0 rows affected(0.00 sec)

mysql□BEGIN;

Query OK,0 rows affected(0.00 sec)

mysql□INSERT INTO t SELECT 1;
```

```
Query OK,1 row affected(0.00 sec)

Records:1 Duplicates:0 Warnings:0

mysql□INSERT INTO t SELECT 1;

ERROR 1062(23000):Duplicate entry'1'for key'PRIMARY'

mysql□SELECT*FROM t\G;

***************************1.row***************************

a:1

1 row in set(0.00 sec)
```

□□□□□□□□□□□□1□□□□□□□□□□□□□1062□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□COMMIT□ROLLBACK□□□

□□□□□□□□□□□ROLLBACK TO SAVEPOINT□□□□ROLLBACK□□□□□□□□□□□□□□□□□□□□□□□□ROLLBACK TO SAVEPOINT□□□□□□□□□□□□COMMIT□ROLLBACK□□□

```
mysql□CREATE TABLE t(a INT,PRIMARY KEY(a))ENGINE=INNODB;

Query OK,0 rows affected(0.00 sec)

mysql□BEGIN;

Query OK,0 rows affected(0.00 sec)

mysql□INSERT INTO t SELECT 1;

Query OK,1 row affected(0.00 sec)

Records:1 Duplicates:0 Warnings:0

mysql□SAVEPOINT t1;

Query OK,0 rows affected(0.00 sec)

mysql□INSERT INTO t SELECT 2;

Query OK,1 row affected(0.00 sec)

Records:1 Duplicates:0 Warnings:0

mysql□SAVEPOINT t2;

Query OK,0 rows affected(0.00 sec)

mysql□RELEASE SAVEPOINT t1;
```

```
Query OK,0 rows affected(0.00 sec)

mysql□INSERT INTO t SELECT 2;

ERROR 1062(23000):Duplicate entry'2'for key'PRIMARY'

mysql□ROLLBACK TO SAVEPOINT t2;

Query OK,0 rows affected(0.00 sec)

mysql□SELECT*FROM t;

+---+

|a|

+---+

|1|

|2|

+---+

2 rows in set(0.00 sec)

mysql□ROLLBACK;

Query OK,0 rows affected(0.00 sec)

mysql□SELECT*FROM t;

Empty set(0.00 sec)
```

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ROLLBACK TO SAVEPOINT t2□□□□□□□□□□t2□□□□□□□□□□□□□□□□□□ROLLBACK□□□□□□□□□□□□□□□□□□□□□□ROLLBACK TO SAVEPOINT□□□□□□□□□□□□□

7.4 □□□□□SQL□□

□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□COMMIT□□□

❑DDL□□□ALTER DATABASE...UPGRADE DATA DIRECTORY NAME□ALTER EVENT□ALTER PROCEDURE□ALTER TABLE□ALTER VIEW□CREATE DATABASE□CREATE EVENT□CREATE INDEX□CREATE PROCEDURE□CREATE TABLE□CREATE TRIGGER□CREATE VIEW□DROP DATABASE□DROP EVENT□DROP INDEX□DROP PROCEDURE□DROP TABLE□DROP TRIGGER□DROP VIEW□RENAME TABLE□TRUNCATE TABLE□

❑□□□□□□□MySQL□□□□□□CREATE USER□DROP USER□GRANT□RENAME USER□REVOKE□SET PASSWORD□

❑□□□□□ANALYZE TABLE□CACHE INDEX□CHECK TABLE□LOAD INDEX INTO CACHE□OPTIMIZE TABLE□REPAIR TABLE□

□□ □□□Microsoft SQL Server□□□□□□□□□□□□□□□□□□DDL□□□□□□□□□□□□□□Microsoft SQL Server□□□□□□□□DDL□□□□□□□□□□InnoDB□□□□□Oracle□□□□□□□□□□

□□□□□□□□□TRUNCATE TABLE□□□DDL□□□□□□□□□□□□DELETE□□□□□□□□□□□□□□□□□□□□□□Microsoft SQL Server□□□□□□□□□

```
mysql□SELECT*FRM t\G;

***************************1.row***************************

a:1

***************************2.row***************************

a:2

2 rows in set(0.00 sec)

mysql□BEGIN;

Query OK,0 rows affected(0.01 sec)
```

```
mysql＞TRUNCATE TABLE t;

Query OK,0 rows affected(0.00 sec)

mysql＞ROLLBACK;

Query OK,0 rows affected(0.00 sec)

mysql＞SELECT*FROM t;

Empty set(0.00 sec)
```

7.5 对于事务操作的统计

由于InnoDB存储引擎是支持事务的，因此InnoDB存储引擎的应用需要在考虑每秒请求数（Question Per Second，QPS）的同时，也应该关注每秒事务处理的能力（Transaction Per Second，TPS）。

计算TPS的方法是（com_commit+com_rollback）/time。但是利用这种方法进行计算的前提是：所有的事务必须都是显式提交的，如果存在隐式地提交和回滚（默认autocommit=1），不会计算到com_commit和com_rollback变量中。如：

```
mysql＞SHOW GLOBAL STATUS LIKE'com_commit'\G;

***************************1.row***************************

Variable_name:Com_commit

Value:5

1 row in set(0.00 sec)

mysql＞INSERT INTO t SELECT 3;

Query OK,1 row affected(0.00 sec)

Records:1 Duplicates:0 Warnings:0

mysql＞SELECT*FROM t\G;

***************************1.row***************************

a:1

***************************2.row***************************

a:2

***************************3.row***************************

a:3

3 rows in set(0.00 sec)

mysql＞SHOW GLOBAL STATUS LIKE'com_commit'\G;

***************************1.row***************************

Variable_name:Com_commit

Value:5
```

```
1 row in set(0.00 sec)
```

MySQL[illegible]handler_commit[illegible]handler_rollback[illegible]MySQL 5.1[illegible]InnoDB[illegible]InnoDB Plugin[illegible]“[illegible]”[illegible]com_commit[illegible]com_rollback[illegible]

7.6 [illegible]

[illegible]ISO[illegible]ANIS SQL[illegible]Oracle[illegible]READ UNCOMMITTED[illegible]REPEATABLE READ[illegible]

SQL[illegible]

❑READ UNCOMMITTED

❑READ COMMITTED

❑REPEATABLE READ

❑SERIALIZABLE

READ UNCOMMITTED[illegible]browse access[illegible]READ COMMITTED[illegible]cursor stability[illegible]REPEATABLE READ[illegible]2.9999°[illegible]SERIALIZABLE[illegible]3°[illegible]SQL[illegible]SQL2[illegible]SERIALIZABLE[illegible]

InnoDB[illegible]REPEATABLE READ[illegible]SQL[illegible]InnoDB[illegible]REPEATABLE READ[illegible]Next-Key Lock[illegible]Microsoft SQL Server[illegible]InnoDB[illegible]REPEATABLE READ[illegible]SQL[illegible]SERIALIZABLE[illegible]

[illegible]READ COMMITTED[illegible]

[illegible]SERIALIZABLE[illegible]Jim Gray[illegible]Transaction Processing[illegible]SERIALIZABLE[illegible]!!![illegible]InnoDB[illegible]REPEATABLE READ[illegible]READ COMMITTED[illegible]

□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
SET[GLOBAL|SESSION]TRANSACTION ISOLATION LEVEL

{

READ UNCOMMITTED

|READ COMMITTED

|REPEATABLE READ

|SERIALIZABLE

}
```

□□□□MySQL□□□□□□□□□□□□□□□□□□□□□□□□□MySQL□□□□□□□[mysqld]□□□□□□□

```
[mysqld]

transaction-isolation=READ-COMMITTED
```

□□□□□□□□□□□□□□□□□□□

```
mysql□SELECT@@tx_isolation\G;

***************************1.row***************************

@@tx_isolation:REPEATABLE-READ

1 row in set(0.01 sec)
```

□□□□□□□□□□□□□□□□□

```
mysql□SELECT@@global.tx_isolation\G;

***************************1.row***************************

@@global.tx_isolation:REPEATABLE-READ

1 row in set(0.00 sec)
```

□SERIALIABLE□□□□□□□□InnoDB□□□□□□□□SELECT□□□□□□□LOCK IN SHARE MODE□□SERIALIZABLE□□□□□□□□□□□□□□□□well-formed□□□□□two-phrased□□□□□□□□□□□□□□□

□□InnoDB□□□□□REPEATABLE READ□□□□□□□□□□3°□□□□□□□□□□□□□□□□□SERIALIABLE□□□□□□SERIALIABLE□□□□□□□□□□□InnoDB□□□□□□□□□□□

□READ COMMITTED□□□□□□□□□□□□□□□□□□□□□□□□□□□□□gap lock□InnoDB□□□□□□□□gap lock□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□MySQL 5.1□□READ COMMITTED□□□□□□□□□□□□□replication□□□□□□□□□□ROW□□□□□□□□□□□□□□□□□□STATEMENT□□□□□□□□□□□□

```
mysql□CREATE TABLE a(

-□b INT,PRIMARY KEY(b)

-□)ENGINE=INNODB;

Query OK,0 rows affected(0.01 sec)

mysql□SET@@tx_isolation='READ-COMMITTED';

Query OK,0 rows affected(0.00 sec)

mysql□SELECT@@tx_isolation\G;

***************************1.row***************************

@@tx_isolation:REPEATABLE-READ

1 row in set(0.00 sec)

mysql□BEGIN;

Query OK,0 rows affected(0.00 sec)

mysql□INSERT INTO a SELECT 1;

ERROR 1598(HY000):Binary logging not possible.Message:Transaction level'READ-COMMITTED'in InnoDB is not safe for binlog
mode'STATEMENT'
```

□MySQL 5.0□□□□□□□□□ROW□□□□□□□□□□□□□□□□□□□□□ innodb_locks_unsafe_for_binlog□□□□1□□□□□□□□□STATEMENT□ □□READ COMMITTED□□□□□□□□

```
mysql□SELCT@@version\G

***************************1.row***************************

@@version:5.0.77-log

1 row in set(0.00 sec)

mysql□SHOW VARIABLES LIKE'innodb_locks_unsafe_for_binlog'\G;

***************************1.row***************************

Variable_name:innodb_locks_unsafe_for_binlog

Value:ON

1 row in set(0.00 sec)

mysql□SET@@tx_isolation='READ-COMMITTED';

Query OK,0 rows affected(0.00 sec)

mysql□BEGIN;

Query OK,0 rows affected(0.00 sec)

mysql□INSERT INTO a SELECT 1;

Query OK,0 rows affected(0.00 sec)

mysql□COMMIT;

Query OK,0 rows affected(0.00 sec)

mysql□SELECT*FROM a\G;

***************************1.row***************************

b:1

***************************2.row***************************

b:2

***************************3.row***************************

b:4

***************************4.row***************************

b:5
```

```
4 rows in set(0.00 sec)
```

□□□master□□□□□□□A□□□□□□□□□□□□□□

```
#Session A on master

mysql□BEGIN;

Query OK,0 rows affected(0.00 sec)

mysql□DELETE FROM a WHERE b□=5;

Query OK,4 rows affected(0.01 sec)
```

□□□□master□□□□□□□□B□□□□□□□□□□□□□

```
#Session B on master

mysql□BEGIN;

Query OK,0 rows affected(0.00 sec)

mysql□INSERT INTO a SELECT 3;

Query OK,0 rows affected(0.01 sec)

mysql□COMMIT;

Query OK,0 rows affected(0.00 sec)
```

□□□□A□□□□□□□a□□□□□

```
#Session A on master

mysql□COMMIT;

Query OK,0 rows affected(0.00 sec)

mysql□SELECT*FROM a\G;

***************************1.row***************************

b:3
```

[illegible]slave[illegible]

```
#Slave

mysql[illegible]SELECT*FROM a;

Empty set(0.00 sec)
```

[illegible]

❑[illegible]READ COMMITTED[illegible]gap lock[illegible]B[illegible]5[illegible]

❑STATEMENT[illegible]master[illegible]SQL[illegible]master[illegible]STATEMENT[illegible]

[illegible]READ REPEATABLE[illegible]master[illegible]slave[illegible]

[illegible]MySQL 5.1[illegible]ROW[illegible]READ COMMITTED[illegible]READ COMMITTED[illegible]ROW[illegible]SQL[illegible]InnoDB[illegible]HeikkiTuuri[illegible]http:‖bugs.mysql.com/bug.php?id=33210[illegible]ROW[illegible]

7.7 [illegible]

7.7.1 MySQL[illegible]

InnoDB[illegible]XA[illegible]XA[illegible]transactional resources[illegible]ACID[illegible]InnoDB[illegible]SERIALIZABLE[illegible]

XA[illegible]MySQL[illegible]Oracle[illegible]SQL Server[illegible]XA[illegible]David[illegible]10 000[illegible]Mariah[illegible]

```
#Bank@Shanghai[illegible]

UPDATE account SET money=money-10000 WHERE user='David';

#Bank@Beijing

UPDATE account SET money=money+10000 WHERE user='Mariah';
```

[illegible]David[illegible]Mariah[illegible]David[illegible]Mariah[illegible]

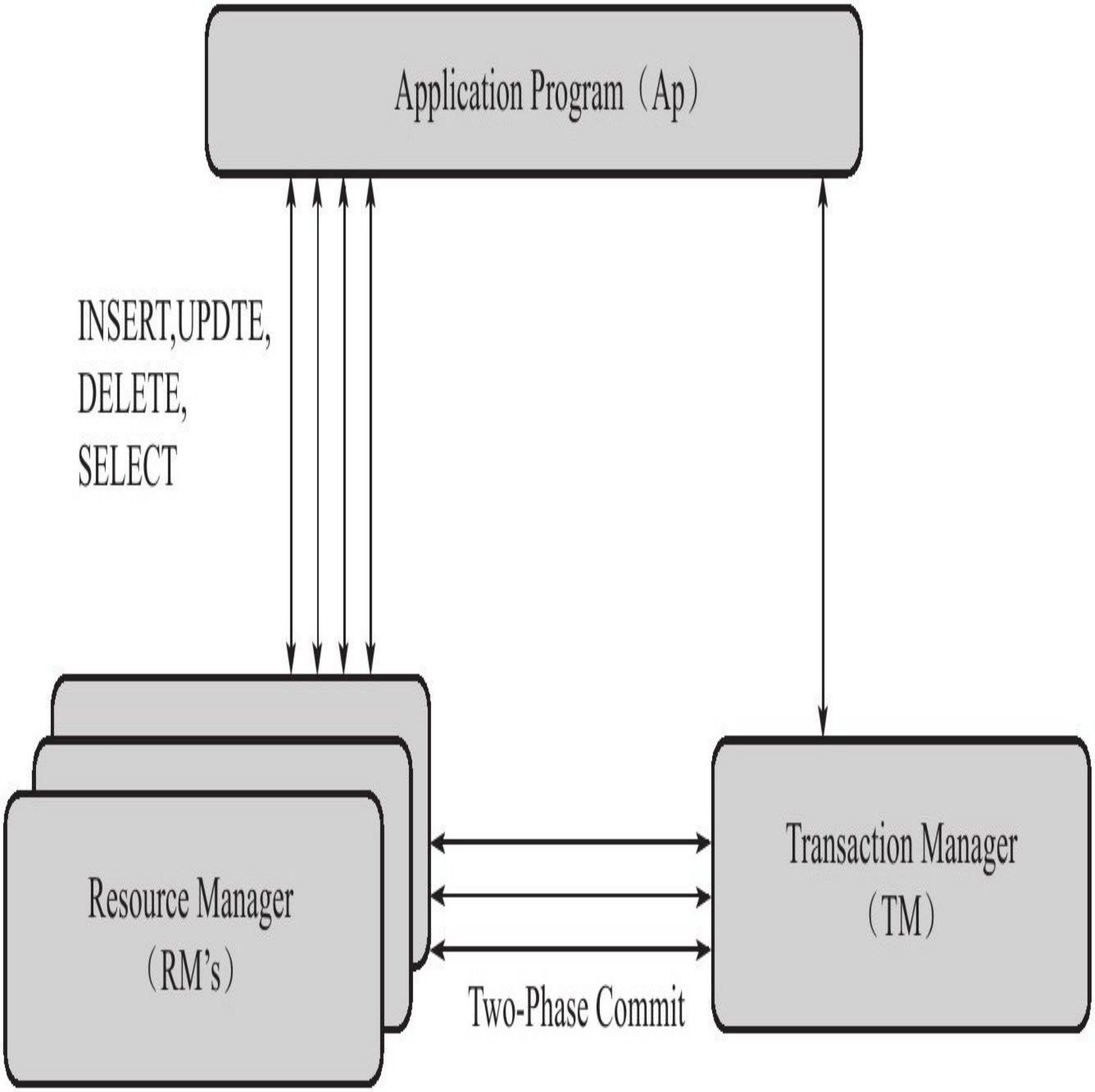
XA[illegible]Resource Managers[illegible]Transaction Manager[illegible]Application Program[illegible]

❑[illegible]

❑[illegible]

❑[illegible]

[illegible]MySQL[illegible]MySQL[illegible]MySQL[illegible]7-22[illegible]



[illegible] 7-22 [illegible]

[illegible]two-phase commit[illegible]PREPARE[illegible]ROLLBACK[illegible]COMMIT[illegible]

□□□□□□□□□□□□□□□□□□□□□PREPARE□□□□□□□□□□□□□□□□□□□□
COMMIT□□ROLLBACK□□□

MySQL□□□XA□□□SQL□□□□□

```
XA{START|BEGIN}xid[JOIN|RESUME]

XA END xid[SUSPEND[FOR MIGRATE]]

XA PREPARE xid

XA COMMIT xid[ONE PHASE]

XA ROLLBACK xid

XA RECOVER
```

□□□□□□□□XA□□□□□□

```
mysql□XA START'a';

Query OK,0 rows affected(0.00 sec)

mysql□INSERT INTO z SELECT 11;

Query OK,1 row affected(0.00 sec)

Records:1 Duplicates:0 Warnings:0

mysql□XA END'a';

Query OK,0 rows affected(0.00 sec)

mysql□XA PREPARE'a';

Query OK,0 rows affected(0.05 sec)

mysql□XA RECOVER\G;

***************************1.row***************************

formatID:1

gtrid_length:1

bqual_length:0

data:a

1 row in set(0.00 sec)
```

```
mysql[illegible]XA COMMIT'a';
Query OK,0 rows affected(0.05 sec)
```

[illegible]MySQL[illegible]Java[illegible]JTA[illegible]Java Transaction API[illegible]MySQL[illegible]API[illegible]JTA[illegible]MySQL[illegible]

```
import java.sql.Connection;
import javax.sql.XAConnection;
import javax.transaction.xa.*;
import com.mysql.jdbc.jdbc2.optional.MysqlXADataSource;
import java.sql.*;
class MyXid implements Xid
{
public int formatId;
public byte gtrid[];
public byte bqual[];
public MyXid(){
}
public MyXid(int formatId,byte gtrid[],byte bqual[])
{
this.formatId=formatId;
this.gtrid=gtrid;
this.bqual=bqual;
}
public int getFormatId()
{
return formatId;
}
```

```
public byte[]getBranchQualifier()

{

return bqual;

}

public byte[]getGlobalTransactionId()

{

return gtrid;

}

}

public class xa_demo{

public static MysqlXADataSource GetDataSource(

String connString,

String user,

String passwd){

try{

MysqlXADataSource ds=new MysqlXADataSource();

ds.setUrl(connString);

ds.setUser(user);

ds.setPassword(passwd);

return ds;

}

catch(Exception e){

System.out.println(e.toString());

return null;

}

}

public static void main(String[]args){

String connString1="jdbc:mysql://192.168.24.43:3306/bank_shanghai";

String connString2="jdbc:mysql://192.168.24.166:3306/bank_

beijing";
```

```
try{
MysqlXADataSource ds1=
GetDataSource(connString1,"peter","12345");
MysqlXADataSource ds2=
GetDataSource(connString2,"david","12345");
XAConnection xaConn1=ds1.getXAConnection();
XAResource xaRes1=xaConn1.getXAResource();
Connection conn1=xaConn1.getConnection();
Statement stmt1=conn1.createStatement();
XAConnection xaConn2=ds2.getXAConnection();
XAResource xaRes2=xaConn2.getXAResource();
Connection conn2=xaConn2.getConnection();
Statement stmt2=conn2.createStatement();
Xid xid1=new MyXid(
100,
new byte[]{0x01},
new byte[]{0x02});
Xid xid2=new MyXid(
100,
new byte[]{0x11},
new byte[]{0x12});
try{
xaRes1.start(xid1,XAResource.TMNOFLAGS);
stmt1.execute("
UPDATE account SET money=money-10000
WHERE user='david'"
);
xaRes1.end(xid1,XAResource.TMSUCCESS);
xaRes2.start(xid2,XAResource.TMNOFLAGS);
stmt2.execute("
UPDATE account SET money=money+10000

WHERE user='mariah'"

);

xaRes2.end(xid2,XAResource.TMSUCCESS);

int ret2=xaRes2.prepare(xid2);

int ret1=xaRes1.prepare(xid1);

if(ret1==XAResource.XA_OK

□□ret2==XAResource.XA_OK){

xaRes1.commit(xid1,false);

xaRes2.commit(xid2,false);

}

}catch(Exception e){

e.printStackTrace();

}

}catch(Exception e){

System.out.println(e.toString());

}

}

}
```

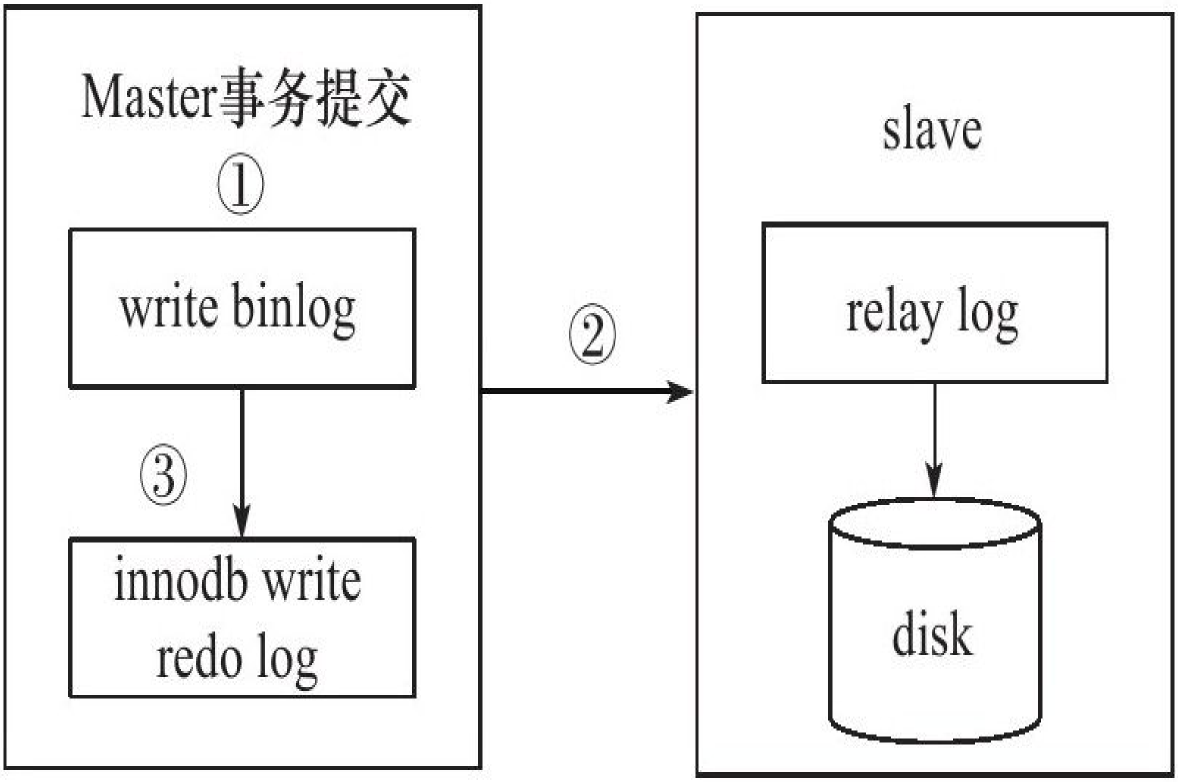
□□□□innodb_support_xa□□□□□□□□□XA□□□□□□□□□ON□□

```
mysql□SHOW VARIABLES LIKE'innodb_support_xa'\G;

***************************1.row***************************

Variable_name:innodb_support_xa

Value:ON

1 row in set(0.01 sec)
```

7.7.2 □□XA□□

□□□□□□□□□□□□□□□□□□□□□□□MySQL□□□□□□□MySQL□□XA□□□

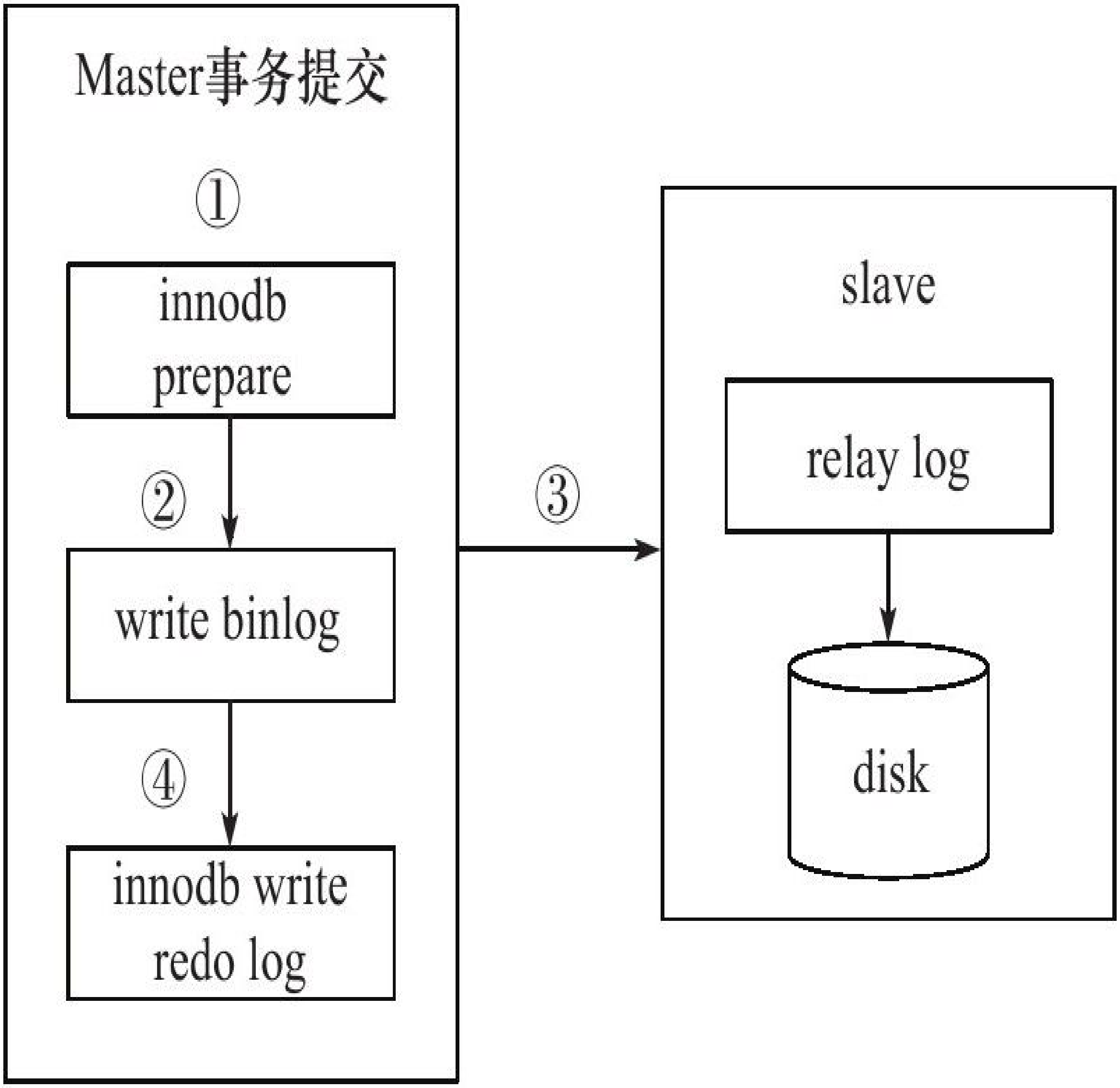
□□□□□□□XA□□□□□binlog□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□binlog□□□□□□□□□□□□□□□□□□□□InnoDB□□InnoDB□□□□□□□□□□□□□slave□□□□□□master□□□□□□□□□□□□□□□□□□□□□□□□□□□□□7-23□□□



□ 7-23 □□□□replication□□□□□□□□

□□7-23□□□□□□□①□②□□□□③□□MySQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□MySQL□□□□binlog□InnoDB□□□□□□□□XA□□□□□□□□□□InnoDB□□□□□□□□□PREPARE□□□□□□□xid□□□□□□□□□□□□□□□□□

[illegible]7-24[illegible]InnoDB[illegible]MySQL[illegible]MySQL[illegible]UXID[illegible]



[illegible] 7-24 MySQL[illegible]XA[illegible]

7.8 [illegible]

7.8.1 [illegible]

[illegible]

```
CREATE PROCEDURE load1(count INT UNSIGNED)

BEGIN

DECLARE s INT UNSIGNED DEFAULT 1;

DECLARE c CHAR(80)DEFAULT REPEAT('a',80);

WHILE s[illegible]=count DO

INSERT INTO t1 SELECT NULL,c;

COMMIT;

SET s=s+1;

END WHILE;

END;
```

[illegible]COMMIT[illegible]InnoDB[illegible]COMMIT[illegible]

```
CREATE PROCEDURE load2(count INT UNSIGNED)

BEGIN

DECLARE s INT UNSIGNED DEFAULT 1;

DECLARE c CHAR(80)DEFAULT REPEAT('a',80);

WHILE s[illegible]=count DO

INSERT INTO t1 SELECT NULL,c;

SET s=s+1;

END WHILE;
```

```
END;
```

[illegible]10 000[illegible]5000[illegible]5000[illegible]load3[illegible]INSERT[illegible]

```
CREATE PROCEDURE load3(count INT UNSIGNED)

BEGIN

DECLARE s INT UNSIGNED DEFAULT 1;

DECLARE c CHAR(80)DEFAULT REPEAT('a',80);

START TRANSACTION;

WHILE s[illegible]=count DO

INSERT INTO t1 SELECT NULL,c;

SET s=s+1;

END WHILE;

COMMIT;

END;
```

[illegible]3[illegible]

```
mysql[illegible]CALL load1(10000);

Query OK,0 rows affected(1 min 3.15 sec)

mysql[illegible]TRUNCATE TABLE t1;

Query OK,0 rows affected(0.05 sec)

mysql[illegible]CALL load2(10000);

Query OK,1 row affected(1 min 1.69 sec)

mysql[illegible]TRUNCATE TABLE t1;

Query OK,0 rows affected(0.05 sec)

mysql[illegible]CALL load3(10000);
```

```
Query OK,0 rows affected(0.63 sec)
```

[illegible]![illegible]load1[illegible]load2[illegible]10 000[illegible]load3[illegible]1[illegible]load2[illegible]load3[illegible]

```
mysql＞BEGIN;

Query OK,0 rows affected(0.00 sec)

mysql＞CALL load2(10000);

Query OK,1 row affected(0.56 sec)

mysql＞COMMIT;

Query OK,0 rows affected(0.03 sec)
```

[illegible]InnoDB[illegible]Oracle[illegible]undo[illegible]Snapshot Too Old[illegible]MySQL[illegible]InnoDB[illegible]

7.8.2 □□□□□□

□□□□□□□□□□□□□□□□□□□□□DBA□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□MySQL□□□□□□□□□□□□□□autocommit□□□□□□□□□□□□□□□□□□□□□□□

```
mysql□SET autocommit=0;

Query OK,0 rows affected(0.00 sec)
```

□□□□□START TRANSACTION□BEGIN□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□completion_type□□0□□MySQL□□□□□□SET AUTOCOMMIT=0□□□□□□COMMIT□ROLLBACK□□□□□□□□□SET AUTOCOMMIT=1□

□□□□□□□□□□API□□□□□□□□□□MySQL C API□□□□□□□□□□□□□□MySQL Python API□□□□□□SET AUTOCOMMIT=0□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□MySQL□API□□□□□

□□□

7.8.3 [illegible]

InnoDB[illegible]HANDLER[illegible]

```
CREATE PROCEDURE sp_auto_rollback_demo()

BEGIN

DECLARE EXIT HANDLER FOR SQLEXCEPTION ROLLBACK;

START TRANSACTION;

INSERT INTO b SELECT 1;

INSERT INTO b SELECT 2;

INSERT INTO b SELECT 1;

INSERT INTO b SELECT 3;

COMMIT;

END;
```

[illegible]sp_auto_rollback_demo[illegible]exit[illegible]HANDLER[illegible]

```
mysql[illegible]SHOW CREATE TABLE b\G;

***************************1.row***************************

Table:b

Create Table:CREATE TABLE'b'(

'a'int(11)NOT NULL DEFAULT'0',

PRIMARY KEY('a')

)ENGINE=InnoDB DEFAULT CHARSET=latin1

1 row in set(0.00 sec)
```

□□□□□□□□□1□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
mysql□CALL sp_auto_rollback_demo;

Query OK,0 rows affected(0.06 sec)

mysql□SELECT*FROM b;

Empty set(0.00 sec)
```

□□□□□□□□□□□□□□□□□□□□sp_auto_rollback_demo□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□sp_auto_rollback_demo□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
CREATE PROCEDURE sp_auto_rollback_demo()

BEGIN

DECLARE EXIT HANDLER FOR SQLEXCEPTION BEGIN ROLLBACK;SELECT-1;END;

START TRANSACTION;

INSERT INTO b SELECT 1;

INSERT INTO b SELECT 2;

INSERT INTO b SELECT 1;

INSERT INTO b SELECT 3;

COMMIT;

SELECT 1;

END;
```

□□□□□□□□□□□□□□-1□□□□□□□□□□□□□□□□1□□□□□□□□□□□□□□□

```
mysql□CALL sp_auto_rollback_demo()\G;

***************************1.row***************************

-1:-1

1 row in set(0.04 sec)

mysql□SELECT*FROM b;
```

```
Empty set(0.00 sec)
```

[illegible]

[illegible]Microsoft SQL Server[illegible] Microsoft SQL Server[illegible]SET XABORT ON[illegible]Microsoft SQL Server[illegible]Microsoft SQL Server[illegible]MySQL[illegible]

[illegible]BEGIN[illegible]COMMIT[illegible]ROLLBACK[illegible]Python[illegible]sp_rollback_demo[illegible]sp_rollback_demo[illegible] sp_auto_rollback_demo[illegible]

```
CREATE PROCEDURE sp_rollback_demo()

BEGIN

INSERT INTO b SELECT 1;

INSERT INTO b SELECT 2;

INSERT INTO b SELECT 1;

INSERT INTO b SELECT 3;

END;
```

[illegible]sp_auto_rollback_demo[illegible]sp_rollback_demo[illegible]test_demo.py[illegible]

```
#!/usr/bin/env python

#encoding=utf-8

import MySQLdb

try:

conn=

MySQLdb.connect(host="192.168.8.7",user="root",passwd="xx",db="test")
```

```
cur=conn.cursor()

cur.execute("SET autocommit=0")

cur.execute("CALL sp_rollback_demo")

cur.execute("COMMIT")

except Exception,e:

cur.execute("ROLLBACK")

print e
```

[illegible]test_demo.py[illegible]

```
[root@nineyou0-43[illegible]]#python test_demo.py

starting rollback

(1062,"Duplicate entry'1'for key'PRIMARY'")
```

[illegible] 1062[illegible]Duplicate entry'1'for key'PRIMARY'[illegible]

7.9 [illegible]

[illegible](Long-Lived Transactions)[illegible]1[illegible]account[illegible]

```
UPDATE account

SET account_total=account_total+(1+interest_rate)
```

[illegible]1[illegible]4[illegible]5[illegible]DBA[illegible]ACID[illegible](mini batch)[illegible]

[illegible]

```
void ComputeInterest[illegible]double interest_rate){

long last_account_done,max_account_no,log_size;

int batch_size=100000;

EXEC SQL SELECT COUNT(*)INTO log_size FROM batchcontext;

if(SQLCODE!=0||log_size==0){

EXEC SQL DROP TABLE IF EXISTS batchcontext;

EXEC SQL CREATE TABLE batchcontext(last_account_done BIGINT);

last_account_done=0;

INSERT INTO batchcontext SELECT 0;

}

else{

EXEC SQL SELECT last_account_no
INTO last_account_done

FROM batchcontext;

}

EXEC SQL SELECT COUNT(*)INTO max_account_no

FROM account LOCK IN SHARE MODE;

WHILE(last_account_no[illegible]max_account_no){

EXEC SQL START TRANSACTION;

EXEC SQL UPDATE account

SET account_total=account_total*(1+interest_rate);

WHERE account_no

BETWEEN last_account_no

AND last_account_no+batch_size;

EXEC SQL UPDATE batchcontext

SET last_account_done=last_account_done+batch_size;

EXEC SQL COMMIT WORK;

last_account_done=last_account_done+batch_size;

}

}
```

[illegible]1[illegible]10[illegible]batchcontext[illegible]ID[illegible]ID[illegible]batchcontext[illegible]1[illegible]batchcontext[illegible]ID[illegible]4000[illegible]40%[illegible]

[illegible]account[illegible]max_account_no[illegible]

7.10 □□

□□□□□□□□□□InnoDB□□

□□□□□□ACID□□□□Atomicity(□□□)□Consistency(□□□)□
Isolation(□□□)□Durability(□□□)□□□□□□□6□□□□□□□□□□□□□□□□□□□□□□□redo□undo□□□□□□□redo□undo□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□REPEATABLE READ□□□□□SQL□□□□□□□□□□□□□□
InnoDB□□□□□REPEATABLE READ□□□□□□□□□□3°□□□□□□

□□□□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□□MySQL□□□□□□□□□□——□□START TRANSACTION□COMMIT□ROLLBACK□□□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□
MySQL InnoDB□□□□□□□

[illegible]8[illegible] [illegible]

[illegible]DBA[illegible]RAID[illegible]DBA[illegible]

[illegible]InnoDB[illegible]MySQL[illegible]mysqldump[illegible]ibbackup[illegible]replication[illegible]xtrabacup[illegible]LVM[illegible]DBA[illegible]

8.1 [illegible]

[illegible]

- Hot Backup[illegible]
- Cold Backup[illegible]
- Warm Backup[illegible]

Hot Backup[illegible]MySQL[illegible]Online Backup[illegible]Cold Backup[illegible]MySQL[illegible]Offline Backup[illegible]Warm Backup[illegible]

[illegible]

- [illegible]
- [illegible]

[illegible]MySQL[illegible]SQL[illegible]mysqldump[illegible]SELECT*INTO OUTFILE[illegible]

[illegible]ibbackup[illegible]xtrabackup[illegible]

[illegible]

- [illegible]
- [illegible]

❑[illegible]

[illegible]MySQL[illegible]replay[illegible]point-in-time[illegible]MySQL[illegible]replication[illegible]slave/standby[illegible]

[illegible]MySQL[illegible]100GB[illegible]SQL[illegible]LSN[illegible]LSN[illegible]xtrabackup[illegible]

[illegible]

[illegible]InnoDB[illegible]MVCC[illegible]REPEATABLE READ[illegible]

[illegible]mysqldump[illegible]--single-transaction[illegible]InnoDB[illegible]InnoDB[illegible]--single-transaction[illegible]mysqldump[illegible]

[illegible]DBA[illegible]24×7[illegible]DAO[illegible]Database Admin Online[illegible]DBA[illegible]Python[illegible]Web[illegible]Django[illegible]DBA[illegible]MySQL[illegible]DBA[illegible]DBA[illegible]

[illegible]2008[illegible]

8.2 [illegible]

[illegible]InnoDB[illegible]MySQL[illegible]frm[illegible]*.ibd[illegible]MySQL[illegible]my.cnf[illegible]

[illegible]DBA[illegible]DBA[illegible]DBA[illegible]

[illegible]

[illegible]

❑[illegible]

❑[illegible]MySQL[illegible]

❑[illegible]

❑[illegible]SQL[illegible]

[illegible]

❑InnoDB[illegible]undo[illegible]

❑[illegible]MySQL[illegible]

8.3 □□□□

8.3.1 mysqldump

mysqldump□□□□□□□Igor Romanenko□□□□□□□□□□□□□□dump□□□□□□□□□□□□□□□□□□□□□MySQL□□□□□□□□□MySQL□□□□□□□□□□□MySQL□□□□□□Oracle□Microsoft SQL Server□□□□□

mysqldump□□□□□□

```
shell□mysqldump[arguments]□fle_name
```

□□□□□□□□□□□□□□□□□--all-databases□□□

```
shell□mysqldump--all-databases□dump.sql
```

□□□□□□□□□□□□□□□□□--databases□□:

```
shell□mysqldump--databases db1 db2 db3□dump.sql
```

□□□□□test□□□□□□□□□□□□□□□□□□

```
[root@xen-server□]#mysqldump--single-transaction test□test_backup.sql
```

□□□□□□□□□□test□□□□□□□□--single-transaction□□□□□□□□□□□□□□□test_backup.sql□□□□□□□□□□□□□□cat□□□□□□□□□□□

```
[root@xen-server□]#cattest_backup.sql

--MySQL dump 10.13 Distrib 5.5.1-m2,for unknown-linux-gnu(x86_64)

--
```

```
--Host:localhost Database:test

-------------------------------------------------------

--Server version 5.5.1-m2-log

……

--

--Table structure for table'a'

--

DROP TABLE IF EXISTS'a';

/*!40101 SET@saved_cs_client=@@character_set_client*/;

/*!40101 SET character_set_client=utf8*/;

CREATE TABLE'a'(

'b'int(11)NOT NULL DEFAULT'0',

PRIMARY KEY('b')

)ENGINE=InnoDB DEFAULT CHARSET=latin1;

/*!40101 SET character_set_client=@saved_cs_client*/;

--

--Dumping data for table'a'

--

LOCK TABLES'a'WRITE;

/*!40000 ALTER TABLE'a'DISABLE KEYS*/;

INSERT INTO'a'VALUES(1),(2),(4),(5);

/*!40000 ALTER TABLE'a'ENABLE KEYS*/;

UNLOCK TABLES;

--

--Table structure for table'z'

--

DROP TABLE IF EXISTS'z';

/*!40101 SET@saved_cs_client=@@character_set_client*/;

/*!40101 SET character_set_client=utf8*/;

CREATE TABLE'z'(
'a'int(11)DEFAULT NULL

)ENGINE=InnoDB DEFAULT CHARSET=latin1;

/*!40101 SET character_set_client=@saved_cs_client*/;

--

--Dumping data for table'z'

--

LOCK TABLES'z'WRITE;

/*!40000 ALTER TABLE'z'DISABLE KEYS*/;

INSERT INTO'z'VALUES(1),(1);

/*!40000 ALTER TABLE'z'ENABLE KEYS*/;

UNLOCK TABLES;

……

--Dump completed on 2010-08-03 13:36:17
```

[illegible]SQL[illegible]MySQL[illegible]CREATE TABLE[illegible]INSERT[illegible]SQL[illegible]

mysqldump[illegible]mysqldump--help[illegible]--lock-tables[illegible]-l[illegible]

❑--single-transaction[illegible]START TRANSACTION[illegible]InnoDB[illegible]DDL[illegible]DDL[illegible]

❑--lock-tables[illegible]-l[illegible]MyISAM[illegible]InnoDB[illegible]--single-transaction[illegible]--lock-tables[illegible]--single-transaction[illegible]exclusive[illegible]MySQL[illegible]MyISAM[illegible]InnoDB[illegible]--lock-tables[illegible]--lock-tables[illegible]

❑--lock-all-tables□-x□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□--lock-tables□□□□□□□□□□□□□□□

❑--add-drop-database□□CREATE DATABASE□□□□DROP DATABASE□□□□□□□□--all-databases□□--databases□□□□□□□□□□□□□□□□□□□□□□□□□□CREATE DATABASE□□□□□□□□□□□□□□□□□□□□□□□□

```
[root@xen-server□]#mysqldump--single-transaction--add-drop-database--databases test□test_backup.sql

[root@xen-server□]#cat test_backup.sql

--MySQL dump 10.13 Distrib 5.5.1-m2,for unknown-linux-gnu(x86_64)

……

--

--Current Database:'test'

--

/*!40000 DROP DATABASE IF EXISTS'test'*/;

CREATE DATABASE/*!32312 IF NOT EXISTS*/'test'/*!40100 DEFAULT CHARACTER SET latin1*/;

USE'test';

……
```

❑--master-data[=value]□□□□□□□□□□□□□□□□□□□□□□□replication□□value□□□1□□□□□□□□□CHANGE MASTER□□□□value□□□2□□CHANGE MASTER□□□□□SQL□□□□□□□□□□value□□□□□□value□□1□□□□□□□□□□□□

```
[root@xen-server□]#mysqldump--single-transaction--add-drop-database--master-data=1--databases test□test_backup.sql

[root@xen-server□]#cat test_backup.sql

--MySQL dump 10.13 Distrib 5.5.1-m2,for unknown-linux-gnu(x86_64)

--

--Host:localhost Database:test

--------------------------------------------------------
```

```
--Server version 5.5.1-m2-log

……

--

--Position to start replication or point-in-time recovery from

--

CHANGE MASTER TO MASTER_LOG_FILE='xen-server-bin.000006',MASTER_LOG_POS=8095;

……
```

□value□2□□□□□□□□□□CHANGE MASTER□□□□□□

```
[root@xen-server□]#mysqldump--single-transaction--add-drop-database--master-data=2--databases test□test_backup.sql

[root@xen-server□]#cat test_backup.sql

--MySQL dump 10.13 Distrib 5.5.1-m2,for unknown-linux-gnu(x86_64)

--

--Host:localhost Database:test

-------------------------------------------------------

--Server version 5.5.1-m2-log

……

--

--Position to start replication or point-in-time recovery from

--

--

……
```

❑--master-data□□□□□--lock-tables□□□□□□□□□--single-transaction□□□□□□□□□--lock-all-tables□□□

❑--events□-E□□□□□□□□□□

❑--routines□-R□□□□□□□□□□□□

❏--triggers□□□□□□□

❏--hex-blob□□BINARY□VARBINARY□BLOG□BIT□□□□□□□□□□□□□□mysqldump□□--hex-blob□□□□□□□□□□□□□□□□□□□

```
[root@xen-server□]#mysqldump--single-transaction--add-drop-database--master-data=2--no-autocommit--databases test3□
test3_backup.sql

[root@xen-server□]#cat test3_backup.sql

--MySQL dump 10.13 Distrib 5.5.1-m2,for unknown-linux-gnu(x86_64)

--

--Host:localhost Database:test3

--------------------------------------------------------

--Server version 5.5.1-m2-log

……

LOCK TABLES'a'WRITE;

/*!40000 ALTER TABLE'a'DISABLE KEYS*/;

setautocommit=0;

INSERT INTO'a'VALUES(0x61000000000000000000);

/*!40000 ALTER TABLE'a'ENABLE KEYS*/;

UNLOCK TABLES;
```

□□□□□□□□0x61000000000000000000□□□□□□□□□□□□□□

❏--tab=path□-T path□□□□TAB□□□□□□□□□□□□□□mysqldump□□□□□□CREATE TABLE□□□table_name.sql□□□□□□□□□tbl_name.txt□□□□□□□--fields-terminated-by=...□--fields-enclosed-by=...□--fields-optionally-enclosed-by=...□--fields-escaped-by=...□--lines-terminated-by=...□□□□□□□□□□□□□□□□□

```
[root@xen-server test]#mysqldump--single-transaction--add-drop-database--tab="/usr/local/mysql/data/test"test
```

```
[root@xen-server test]#ls-lh
total 244K
-rw-rw----1 mysql mysql 8.4K Jul 21 16:02 a.frm
-rw-rw----1 mysql mysql 96K Jul 22 17:18 a.ibd
-rw-r--r--1 root root 1.3K Aug 3 15:36 a.sql
-rw-rw-rw-1 mysql mysql 8 Aug 3 15:36 a.txt
-rw-rw----1 mysql mysql 65 Jul 17 15:54 db.opt
-rw-rw----1 mysql mysql 8.4K Aug 2 17:22 z.frm
-rw-rw----1 mysql mysql 96K Aug 2 17:22 z.ibd
-rw-r--r--1 root root 1.3K Aug 3 15:36 z.sql
-rw-rw-rw-1 mysql mysql 4 Aug 3 15:36 z.txt
-----------
--Server version 5.5.1-m2-log
/*!40101 SET@OLD_CHARACTER_SET_CLIENT=@@CHARACTER_SET_CLIENT*/;
/*!40101 SET@OLD_CHARACTER_SET_RESULTS=@@CHARACTER_SET_RESULTS*/;
/*!40101 SET@OLD_COLLATION_CONNECTION=@@COLLATION_CONNECTION*/;
/*!40101 SET NAMES utf8*/;
/*!40103 SET@OLD_TIME_ZONE=@@TIME_ZONE*/;
/*!40103 SET TIME_ZONE='+00:00'*/;
/*!40101 SET@OLD_SQL_MODE=@@SQL_MODE,SQL_MODE=''*/;
/*!40111 SET@OLD_SQL_NOTES=@@SQL_NOTES,SQL_NOTES=0*/;
--
--Table structure for table'a'
--
DROP TABLE IF EXISTS'a';
/*!40101 SET@saved_cs_client=@@character_set_client*/;
/*!40101 SET character_set_client=utf8*/;
CREATE TABLE'a'(
'b'int(11)NOT NULL DEFAULT'0',
PRIMARY KEY('b')
)ENGINE=InnoDB DEFAULT CHARSET=latin1;

/*!40101 SET character_set_client=@saved_cs_client*/;

/*!40103 SET TIME_ZONE=@OLD_TIME_ZONE*/;

/*!40101 SET SQL_MODE=@OLD_SQL_MODE*/;

/*!40101 SET CHARACTER_SET_CLIENT=@OLD_CHARACTER_SET_CLIENT*/;

/*!40101 SET CHARACTER_SET_RESULTS=@OLD_CHARACTER_SET_RESULTS*/;

/*!40101 SET COLLATION_CONNECTION=@OLD_COLLATION_CONNECTION*/;

/*!40111 SET SQL_NOTES=@OLD_SQL_NOTES*/;

--Dump completed on 2010-08-03 15:36:56

[root@xen-server test]#cat a.txt

1

2

4

5
```

[illegible]DBA[illegible]SELECT...INTO OUTFILE[illegible] mysqldump[illegible]

❑--where='where_condition'[illegible]-w'where_condition'[illegible]b[illegible]a[illegible]a[illegible]2[illegible]

```
[root@xen-server bin]#mysqldump--single-transaction--where='b□2'test a□a.sql

[root@xen-server bin]#cat a.sql

--MySQL dump 10.13 Distrib 5.5.1-m2,for unknown-linux-gnu(x86_64)

--

--Host:localhost Database:test

--------------------------------------------------------

--Server version 5.5.1-m2-log

……

--

--Dumping data for table'a'
```

```
--
--WHERE:b□2
LOCK TABLES'a'WRITE;
/*!40000 ALTER TABLE'a'DISABLE KEYS*/;
INSERT INTO'a'VALUES(4),(5);
/*!40000 ALTER TABLE'a'ENABLE KEYS*/;
UNLOCK TABLES;
/*!40103 SET TIME_ZONE=@OLD_TIME_ZONE*/;
……
```

8.3.2 SELECT...INTO OUTFILE

SELECT...INTO□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
SELECT...INTO□□□□□□

```
SELECT[column 1],[column 2]...

INTO

OUTFILE'file_name'

[{FIELDS|COLUMNS}

[TERMINATED BY'string']

[[OPTIONALLY]ENCLOSED BY'char']

[ESCAPED BY'char']

]

[LINES

[STARTING BY'string']

[TERMINATED BY'string']

]

FROM TABLE WHERE......
```

□□FIELDS[TERMINATED BY'string']□□□□□□□□□□[[OPTIONALLY]ENCLOSED BY'char']□□□□□□□□□□□□[ESCAPED BY'char']□□□□□□[STARTING BY'string']□□□□□□□□□□TERMINATED BY'string'□□□□□□□□□□□□□□□□□□□FIELDS□LINES□□□□□□□□□□□□□□

```
FIELDS TERMINATED BY'\t'ENCLOSED BY''ESCAPED BY'\\'

LINES TERMINATED BY'\n'STARTING BY''
```

file_name□□□□□□□□□□□□□□□□□□□□□□mysql:mysql□□□□MySQL□□□□□□□□□

```
mysql□select*into outfile'/root/a.txt'from a;

ERROR 1(HY000):Can't create/write to file'/root/a.txt'(Errcode:13)
```

□□□□□□□□□□□□□□□□

```
[root@xen-server□]#mysql test-e"select*into outfile'/home/mysql/a.txt'fields terminated by','from a";

ERROR 1086(HY000)at line 1:File'/home/mysql/a.txt'already exists
```

□□□□SELECT INTO□□□□a□□□

```
mysql□select*into outfile'/home/mysql/a.txt'from a;

Query OK,3 rows affected(0.02 sec)

mysql□quit

Bye

[root@xen-server□]#cat/home/mysql/a.txt

1 a

2 b

3 c
```

□□□□□□□□□□□□□□TAB□□□□□□□□□□□□□□□□□□□□“□”□□□□□□FIELDS TERMINATED BY'string'□□□□□

```
[root@xen-server□]#mysql test-e"select*into outfile'/home/mysql/a.txt'fields terminated by','from a";

[root@xen-server□]#cat/home/mysql/a.txt

1,a

2,b

3,c
```

□Windows□□□□□□□□□□“\r\n”□□□□□□□□□□□□□LINES TERMINATED BY□□□□□

```
[root@xen-servermysql]#mysql test-e"select*into outfile'/home/mysql/a.txt'fields terminated by','lines terminated by'\r\n'from
a";

[root@xen-servermysql]#od-c a.txt

0000000 1,a\r\n 2,b\r\n 3,c\r\n

0000017
```

8.3.3 [illegible]

mysqldump[illegible]SQL[illegible]

```
[root@xen-server□]#mysql-uroot-p□test_backup.sql

Enter password:
```

[illegible]SQL[illegible]

```
mysql□drop database test;

ERROR 1010(HY000):Error dropping database(can't rmdir'./test',errno:39)
```

[illegible]SQL[illegible]SOURCE[illegible]

```
mysql□source/home/mysql/test_backup.sql;

Query OK,0 rows affected(0.00 sec)

Query OK,0 rows affected(0.00 sec)

……

Query OK,0 rows affected(0.00 sec)

Query OK,0 rows affected(0.00 sec)
```

[illegible]mysqldump[illegible]mysqldump[illegible]mysqldump[illegible]frm[illegible]mysqldump[illegible]

8.3.4 LOAD DATA INFILE

[illegible]mysqldump-tab[illegible]SELECT INTO OUTFILE[illegible]LOAD DATA INFILE[illegible]LOAD DATA INFILE[illegible]

```
LOAD DATA INTO TABLE a IGNORE 1 LINES INFILE'/home/mysql/a.txt'
[REPLACE|IGNORE]
INTO TABLE tbl_name
[CHARACTER SET charset_name]
[{FIELDS|COLUMNS}
[TERMINATED BY'string']
[[OPTIONALLY]ENCLOSED BY'char']
[ESCAPED BY'char']
]
[LINES
[STARTING BY'string']
[TERMINATED BY'string']
]
[IGNORE number LINES]
[(col_name_or_user_var,...)]
[SET col_name=expr,...]
```

[illegible]LOAD DATA INFILE[illegible]FILE[illegible]SELECT INTO OUTFILE[illegible]IGNORE number LINES[illegible]LOAD DATA INFILE[illegible]

```
mysql[illegible]load data infile'/home/mysql/a.txt'into table a;
Query OK,3 rows affected(0.00 sec)
Records:3 Deleted:0 Skipped:0 Warnings:0
```

□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
mysql□SET@@foreign_key_checks=0;

Query OK,0 rows affected(0.00 sec)

mysql□LOAD DATA INFILE'/home/mysql/a.txt'INTO TABLE a;

Query OK,4 rows affected(0.00 sec)

Records:4 Deleted:0 Skipped:0 Warnings:0

mysql□SET@@foreign_key_checks=1;

Query OK,0 rows affected(0.00 sec)
```

□□□□□□□□□□□□□□□□□□□□□□a□b□□c□□□a□b□□□□

```
mysql□CREATE TABLE b(

-□a INT,

-□b INT,

-□c INT,

-□PRIMARY KEY(a)

-□)ENGINE=InnoDB;

Query OK,0 rows affected(0.01 sec)

mysql□LOAD DATA INFILE'/home/mysql/a.txt'

-□INTO TABLE b FIELDS TERMINATED BY','(a,b)

-□SET c=a+b;

Query OK,4 rows affected(0.01 sec)

Records:4 Deleted:0 Skipped:0 Warnings:0

mysql□SELECT*FROM b;

+---+------+------+

|a|b|c|

+---+------+------+

|1|2|3|
|2|3|5|

|4|5|9|

|5|6|11|

+---+------+------+

4 rows in set(0.00 sec)
```

8.3.5 mysqlimport

mysqlimport[illegible]MySQL[illegible]LOAD DATA INFILE[illegible]LOAD DATA INFILE[illegible]

```
shell[illegible]mysqlimport[options]db_name textfile1[textfile2...]
```

[illegible]LOAD DATA INFILE[illegible]mysqlimport[illegible]--user-thread[illegible]mysqlimport[illegible]mysqlimport[illegible]2[illegible]

```
[root@xen-servermysql]#mysqlimport--use-threads=2 test/home/mysql/t.txt/home/mysql/s.txt

test.s:Records:5000000 Deleted:0 Skipped:0 Warnings:0

test.t:Records:5000000 Deleted:0 Skipped:0 Warnings:0
```

[illegible]MySQL[illegible]

```
mysql[illegible]SHOW FULL PROCESSLIST\G;

***************************1.row***************************

Id:46

User:rep

Host:www.dao.com:1028

db:NULL

Command:Binlog Dump

Time:37651

State:Master has sent all binlog to slave;waiting for binlog to be updated

Info:NULL

***************************2.row***************************
```

```
Id:77
User:root
Host:localhost
db:test
Command:Query
Time:0
State:NULL
Info:show full processlist
***************************3.row***************************
Id:83
User:root
Host:localhost
db:test
Command:Query
Time:73
State:NULL
Info:LOAD DATA INFILE'/home/mysql/t.txt'INTO TABLE't'IGNORE 0 LINES
***************************4.row***************************
Id:84
User:root
Host:localhost
db:test
Command:Query
Time:73
State:NULL
Info:LOAD DATA INFILE'/home/mysql/s.txt'INTO TABLE's'IGNORE 0 LINES
4 rows in set(0.00 sec)
```

[illegible]mysqlimport[illegible]LOAD DTA INFILE[illegible]

8.4 □□□□□□□□□□

□□□□□□□□□□□□□□□□□□□point-in-time□□□□□□MySQL□□□□replication□□

```
[mysqld]

log-bin=mysql-bin
```

□3.2.4□□□□□□□□□□InnoDB□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□

```
[mysqld]

log-bin=mysql-bin

sync_binlog=1

innodb_support_xa=1
```

□□□□□□□□□□□□□□□□FLUSH LOGS□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□mysqlbinlog□□□mysqlbinlog□□□□□□□□

```
shell□mysqlbinlog[options]log_fle...
```

□□□□□binlog.0000001□□□□□□□□□□

```
shell□mysqlbinlog binlog.0000001|mysql-uroot-p test
```

[illegible]

```
shell[illegible]mysqlbinlog binlog.[0-10]*|mysql-u root-p test
```

[illegible]mysqlbinlog[illegible]SOURCE[illegible]

```
shell[illegible]mysqlbinlog binlog.000001[illegible]/tmp/statements.sql

shell[illegible]mysqlbinlog binlog.000002[illegible]/tmp/statements.sql

shell[illegible]mysql-u root-p-e"source/tmp/statements.sql"
```

--start-position[illegible]--stop-position[illegible]

```
shell[illegible]mysqlbinlog--start-position=107856 binlog.0000001|mysql-uroot-p test
```

--start-datetime[illegible]--stop-datetime[illegible]--start-position[illegible]--stop-position[illegible]

8.5 [illegible]

8.5.1 ibbackup

ibbackup[illegible]InnoDB[illegible]MyISAM[illegible]InnoDB[illegible]InnoDB[illegible]

1[illegible]InnoDB[illegible]LSN[illegible]

2[illegible]

3[illegible]InnoDB[illegible]LSN[illegible]

4[illegible]

[illegible]Microsoft SQL Server[illegible]Oracle[illegible]ibbackup[illegible]

❑[illegible]SQL[illegible]

❑[illegible]

❑[illegible]

❑[illegible]ibbackup[illegible]Linux[illegible]Windows[illegible]UNIX[illegible]

ibbackup[illegible]InnoDB[illegible]

❑[illegible]

❑[illegible]

ibbackup[illegible]InnoDB[illegible]Percona[illegible]

XtraBackup□□□□□□□□□□ibbackup□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□XtraBackup□□□□□□□□□

8.5.2 XtraBackup

XtraBackup□□□□□□Percona□□□□□□□□□□□□□□MySQL5.0□□□□□□XtraBackup□GPL v2□□□□□□□□□□□□https://launchpad.net/percona-xtrabackup□

xtrabackup□□□□□□□□□□

```
xtrabackup--backup|--prepare[OPTIONS]
```

xtrabackup□□□□□□□□□□

```
(The defaults options should be given as the first argument)

--print-defaults Prints the program's argument list and exit.

--no-defaults Don't read the default options from any file.

--defaults-file=Read the default options from this file.

--defaults-extra-file=Read this file after the global options files have been read.

--target-dir=The destination directory for backups.

--backup Make a backup of a mysql instance.

--stats Calculate the statistic of the datadir(it is recommended you take mysqld offline).

--prepare Prepare a backup so you can start mysql server with your restore.

--export Create files to import to another database after it has been prepared.

--print-param Print the parameters of mysqld that you will need for a forcopyback.

--use-memory=This value is used instead of buffer_pool_size.

--suspend-at-end Creates a file called xtrabackup_suspended and waits until the user deletes that file at the end of the
backup.

--throttle=(use with--backup)Limits the IO operations(pairs of reads and writes)per second to the values set here.

--log-stream outputs the contents of the xtrabackup_logfile to stdout.

--incremental-lsn=(use with--backup)Copy only.ibd pages newer than the specified LSN high:low.##ATTENTION##:checkpoint
lsn*must*be used.Be Careful!

--incremental-basedir=(use with--backup)Copy only.ibd pages newer than the existing backup at the specified directory.
```

```
--incremental-dir=(use with--prepare)Apply.delta files and logfiles located in the specified directory.
--tables=name Regular Expression list of table names to be backed up.
--create-ib-logfile(NOT CURRENTLY IMPLEMENTED)will create ib_logfile*after a--prepare.
###If you want to create ib_logfile*only re-execute this
command using the same options.###
--datadir=name Path to the database root.
--tmpdir=name Path for temporary files.Several paths may be specified as a colon(:)separated string.
If you specify multiple paths they are used round-robin.
```

□□□□□□□□□□□□□□□□□□□□□□

```
#./xtrabackup--backup
./xtrabackup Ver alpha-0.2 for 5.0.75 unknown-linux-gnu(x86_64)
□□log scanned up to(0 1009910580)
Copying./ibdata1
to/home/kinoyasu/xtrabackup_work/mysql-5.0.75/innobase/xtrabackup/tmp2/ibdata1
...done
Copying./tpcc/stock.ibd
to/home/kinoyasu/xtrabackup_work/mysql-5.0.75/innobase/xtrabackup/tmp2/tpcc/stock.ibd
...done
Copying./tpcc/new_orders.ibd
to/home/kinoyasu/xtrabackup_work/mysql-5.0.75/innobase/xtrabackup/tmp2/tpcc/new_orders.ibd
...done
Copying./tpcc/history.ibd
to/home/kinoyasu/xtrabackup_work/mysql-5.0.75/innobase/xtrabackup/tmp2/tpcc/history.ibd
...done
Copying./tpcc/customer.ibd
to/home/kinoyasu/xtrabackup_work/mysql-5.0.75/innobase/xtrabackup/tmp2/tpcc/customer.ibd
□□log scanned up to(0 1010561109)
...done
```

```
Copying./tpcc/district.ibd

to/home/kinoyasu/xtrabackup_work/mysql-5.0.75/innobase/xtrabackup/tmp2/tpcc/district.ibd

...done

Copying./tpcc/item.ibd

to/home/kinoyasu/xtrabackup_work/mysql-5.0.75/innobase/xtrabackup/tmp2/tpcc/item.ibd

...done

Copying./tpcc/order_line.ibd

to/home/kinoyasu/xtrabackup_work/mysql-5.0.75/innobase/xtrabackup/tmp2/tpcc/order_line.ibd

□□log scanned up to(0 1012047066)

...done

Copying./tpcc/orders.ibd

to/home/kinoyasu/xtrabackup_work/mysql-5.0.75/innobase/xtrabackup/tmp2/tpcc/orders.ibd

...done

Copying./tpcc/warehouse.ibd

to/home/kinoyasu/xtrabackup_work/mysql-5.0.75/innobase/xtrabackup/tmp2/tpcc/warehouse.ibd

...done

□□log scanned up to(0 1014592707)

Stopping log copying thread..

Transaction log of lsn(0 1009910580)to(0 1014592707)was copied.
```

□□□□□□□□□□□xtrabackup□□□□□□□□□□□□□□□□□□□□□0 1009910580□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□copy□□□□□□□□□□□□Copying…to…□□□□□□□□□□□□□□□□□□0 1014592707□□

8.5.3 XtraBackup[illegible]

MySQL[illegible]point-in-time[illegible]XtraBackup[illegible]InnoDB[illegible]

1[illegible]LSN[illegible]

2[illegible]LSN[illegible]LSN[illegible]LSN[illegible]

[illegible]XtraBackup[illegible]

```
(full backup)

#./xtrabackup--backup--target-dir=/backup/base

...

(incremental backup)

#./xtrabackup--backup--target-dir=/backup/delta--incremental-basedir=/backup/base

...

(prepare)

#./xtrabackup--prepare--target-dir=/backup/base

...

(apply incremental backup)

#./xtrabackup--prepare--target-dir=/backup/base--incremental-dir=/backup/delta

...
```

[illegible]/backup/base[illegible]/backup/delta[illegible]

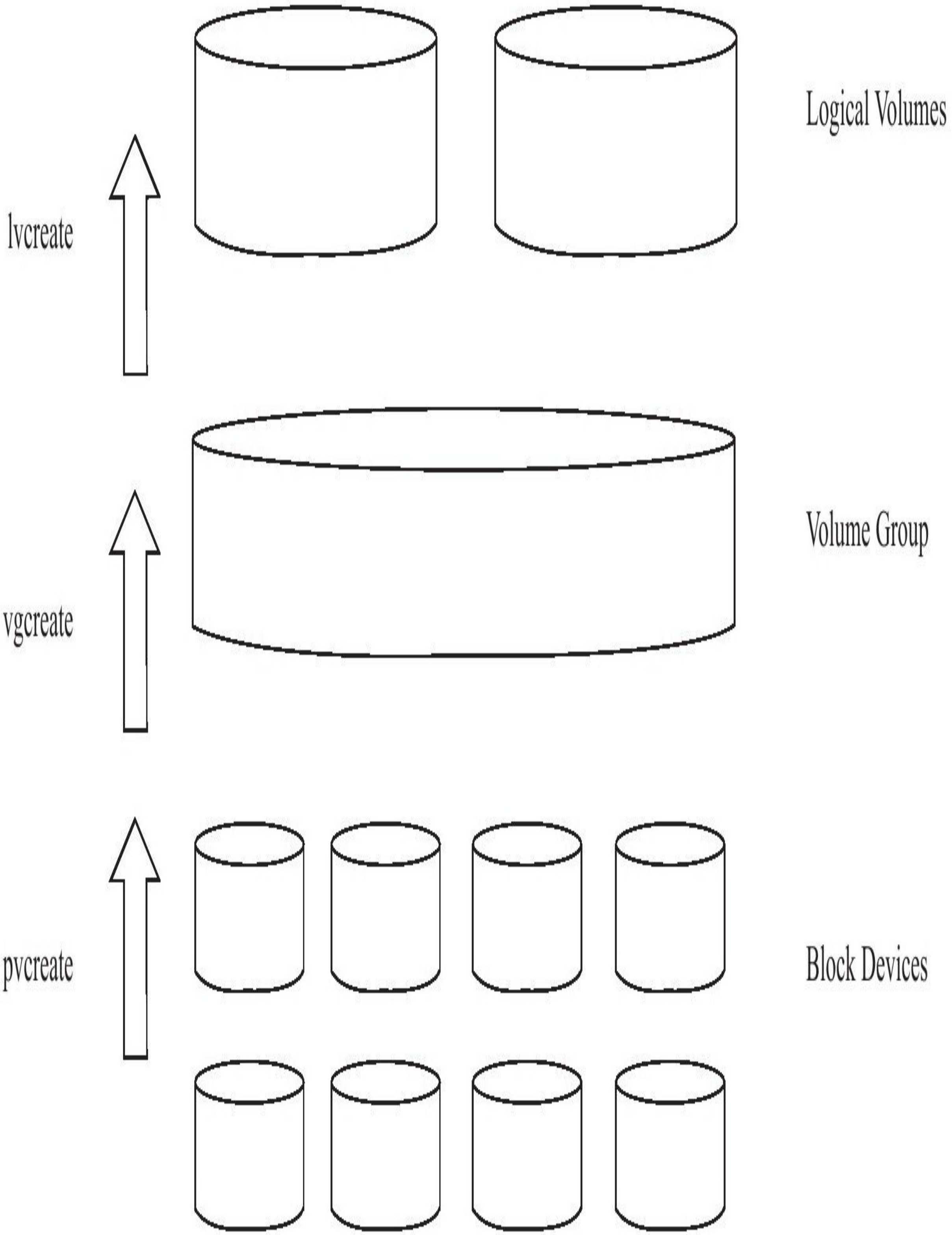
```
#./xtrabackup--backup
```

```
./xtrabackup Ver beta-0.4 for 5.0.75 unknown-linux-gnu(x86_64)

□□log scanned up to(0 378161500)

...

The latest check point(for incremental):'0:377883685'□=====□□□□LSN

□□log scanned up to(0 379294296)

Stopping log copying thread..

Transaction log of lsn(0 377883685)to(0 379294296)was copied.

(must do--prepare before the each incremental backup)

#./xtrabackup--prepare

...

#./xtrabackup--backup--incremental=0:377883685

incremental backup from 0:377883685 is enabled.

./xtrabackup Ver beta-0.4 for 5.0.75 unknown-linux-gnu(x86_64)

□□log scanned up to(0 379708047)

Copying./ibdata1

to/home/kinoyasu/xtrabackup_work/mysql-5.0.75/innobase/xtrabackup/tmp_diff/ibdata1.delta

...done

...

The latest check point(for incremental):'0:379438233'□====□□□□□□□□□LSN

□□log scanned up to(0 380663549)

Stopping log copying thread..

Transaction log of lsn(0 379438233)to(0 380663549)was copied.
```

8.6 [illegible]

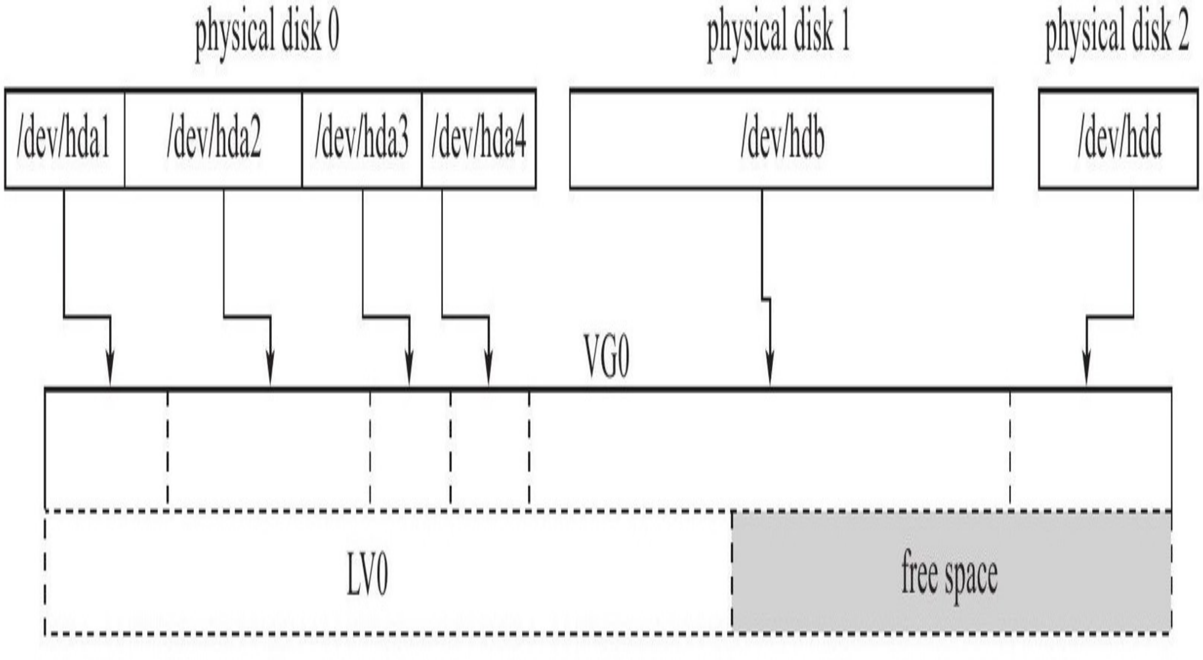
MySQL[illegible]FreeBSD[illegible]UFS[illegible]Solaris[illegible]ZFS[illegible]GNU/Linux[illegible]Logical Volume Manager[illegible]LVM[illegible]LVM[illegible]UFS[illegible]ZFS[illegible]LVM[illegible]

LVM[illegible]LINUX[illegible]LVM[illegible]LVM[illegible]Volume Group[illegible]Logical Volumes[illegible]LVM[illegible]LVM[illegible]8-1[illegible]


Logical Volumes
lvcreate
Volume Group
vgcreate
Block Devices
pvcreate

□ 8-1 LVM□□□□

□8-2□□□□□□□□□□□□□□LV0□



□ 8-2 □□□□□□□□□

□□vgdisplay□□□□□□□□□□□□□□□

```
[root@nh124-98□]#vgdisplay

---Volume group---

VG Name rep

System ID

Format lvm2

Metadata Areas 1

Metadata Sequence No 1873

VG Access read/write

VG Status resizable

MAX LV 0
```

```
Cur LV 3

Open LV 1

Max PV 0

Cur PV 1

Act PV 1

VG Size 260.77 GB

PE Size 4.00 MB

Total PE 66758

Alloc PE/Size 66560/260.00 GB

Free PE/Size 198/792.00 MB

VG UUID MQJiye-j4NN-LbZG-F3CQ-UdTU-fo9D-RRfXD5
```

vgdisplay□□□□□□□□□□□□□□□□rep□□□□□□□260.77GB□□□□□□□□□read/write□□□□lvdisplay□□□□□□□□□□□□□□□□□□

```
[root@nh124-98□]#lvdisplay

---Logical volume---

LV Name/dev/rep/repdata

VG Name rep

LV UUID 7tOlDt-seKZ-ChpY-QMXC-WaFD-zXAl-MRbofK

LV Write Access read/write

LV snapshot status source of

/dev/rep/dho_datasnapshot100805143507[active]

/dev/rep/dho_datasnapshot100805163504[active]

LV Status available

#open 1

LV Size 100.00 GB

Current LE 25600

Segments 1

Allocation inherit
```

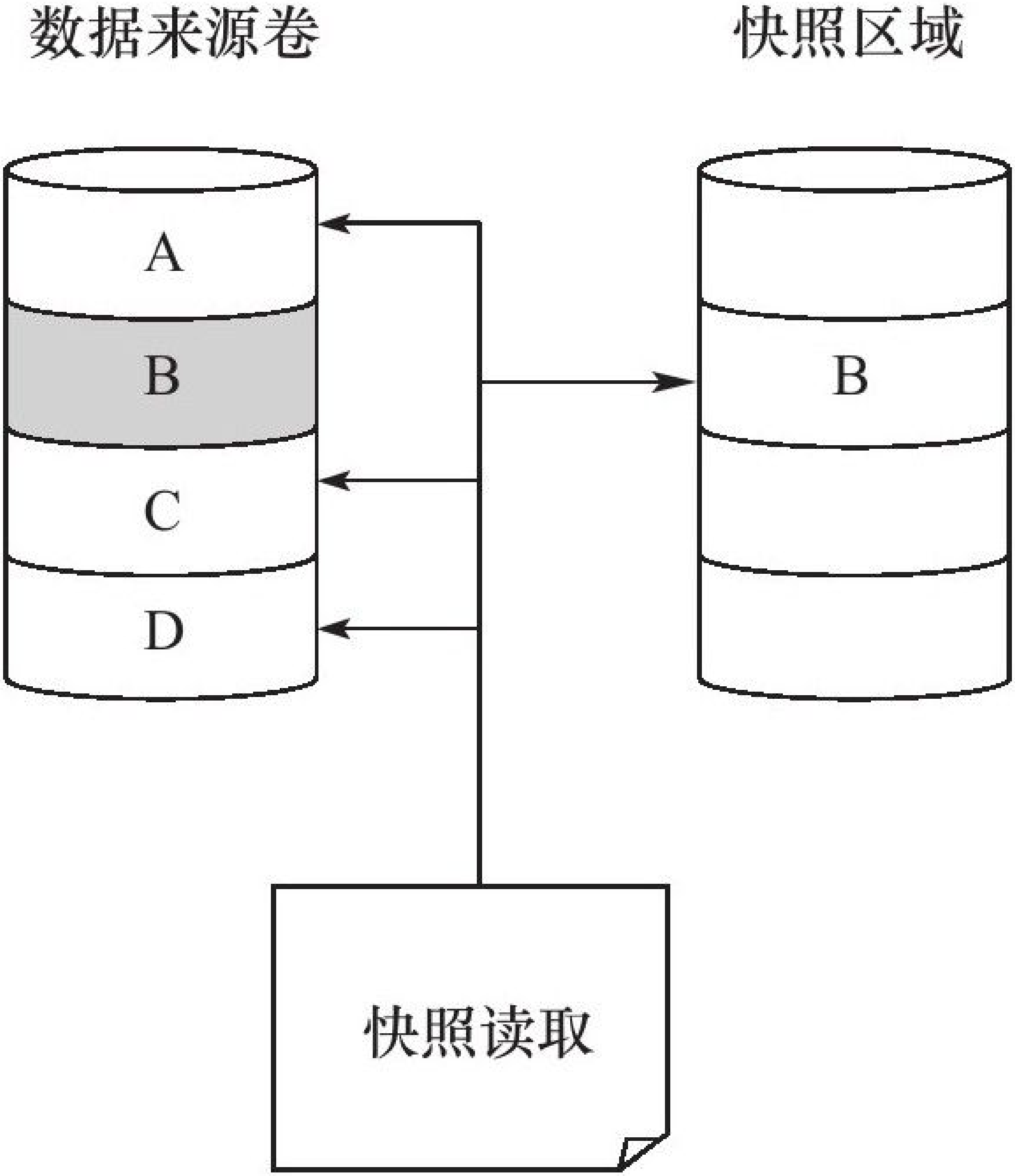
```
Read ahead sectors auto

-currently set to 256

Block device 253:0

---Logical volume---

LV Name/dev/rep/dho_datasnapshot100805143507

VG Name rep

LV UUID fSSXzh-IBnZ-aZIn-eP03-b7pk-CPjN-5xUktE

LV Write Access read only

LV snapshot status active destination for/dev/rep/repdata

LV Status available

#open 0

LV Size 100.00 GB

Current LE 25600

COW-table size 80.00 GB

COW-table LE 20480

Allocated to snapshot 0.13%

Snapshot chunk size 4.00 KB

Segments 1

Allocation inherit

Read ahead sectors auto

-currently set to 256

Block device 253:1

---Logical volume---

LV Name/dev/rep/dho_datasnapshot100805163504

VG Name rep

LV UUID 3B9NP1-qWVG-pfJY-Bdgm-DIdD-dUMu-s2L6qJ

LV Write Access read only

LV snapshot status active destination for/dev/rep/repdata

LV Status available

#open 0
```

```
LV Size 100.00 GB

Current LE 25600

COW-table size 80.00 GB

COW-table LE 20480

Allocated to snapshot 0.02%

Snapshot chunk size 4.00 KB

Segments 1

Allocation inherit

Read ahead sectors auto

-currently set to 256

Block device 253:4
```

[illegible]3[illegible]rep[illegible]100GB[illegible]/dev/rep/repdata[illegible]

LVM[illegible]Copy-on-write[illegible]meta data[illegible]

[illegible]8-3[illegible]LVM[illegible]B[illegible]A[illegible]C[illegible]D[illegible]B[illegible]



[illegible] 8-3 LVM[illegible]

[illegible]lvcreate[illegible]--permission r[illegible]

```
[root@nh119-215 data]#lvcreate--size 100G--snapshot--permission r-n datasnapshot/dev/rep/repdata
```

```
Logical volume"datasnapshot"created
```

[illegible]lvdisplay[illegible]COW-table size[illegible]Allocated to snapshot[illegible]

```
[root@nh124-98[illegible]]#lvdisplay

……

---Logical volume---

LV Name/dev/rep/dho_datasnapshot100805163504

VG Name rep

LV UUID 3B9NP1-qWVG-pfJY-Bdgm-DIdD-dUMu-s2L6qJ

LV Write Access read only

LV snapshot status active destination for/dev/rep/repdata

LV Status available

#open 0

LV Size 100.00 GB

Current LE 25600

COW-table size 80.00 GB

COW-table LE 20480

Allocated to snapshot 0.04%

Snapshot chunk size 4.00 KB

Segments 1

Allocation inherit

Read ahead sectors auto

-currently set to 256

Block device 253:4
```

[illegible]0.04%[illegible]

[illegible]LVM[illegible]InnoDB[illegible]InnoDB[illegible]

[illegible]InnoDB[illegible]InnoDB[illegible]

8.7 [illegible]

8.7.1 [illegible]

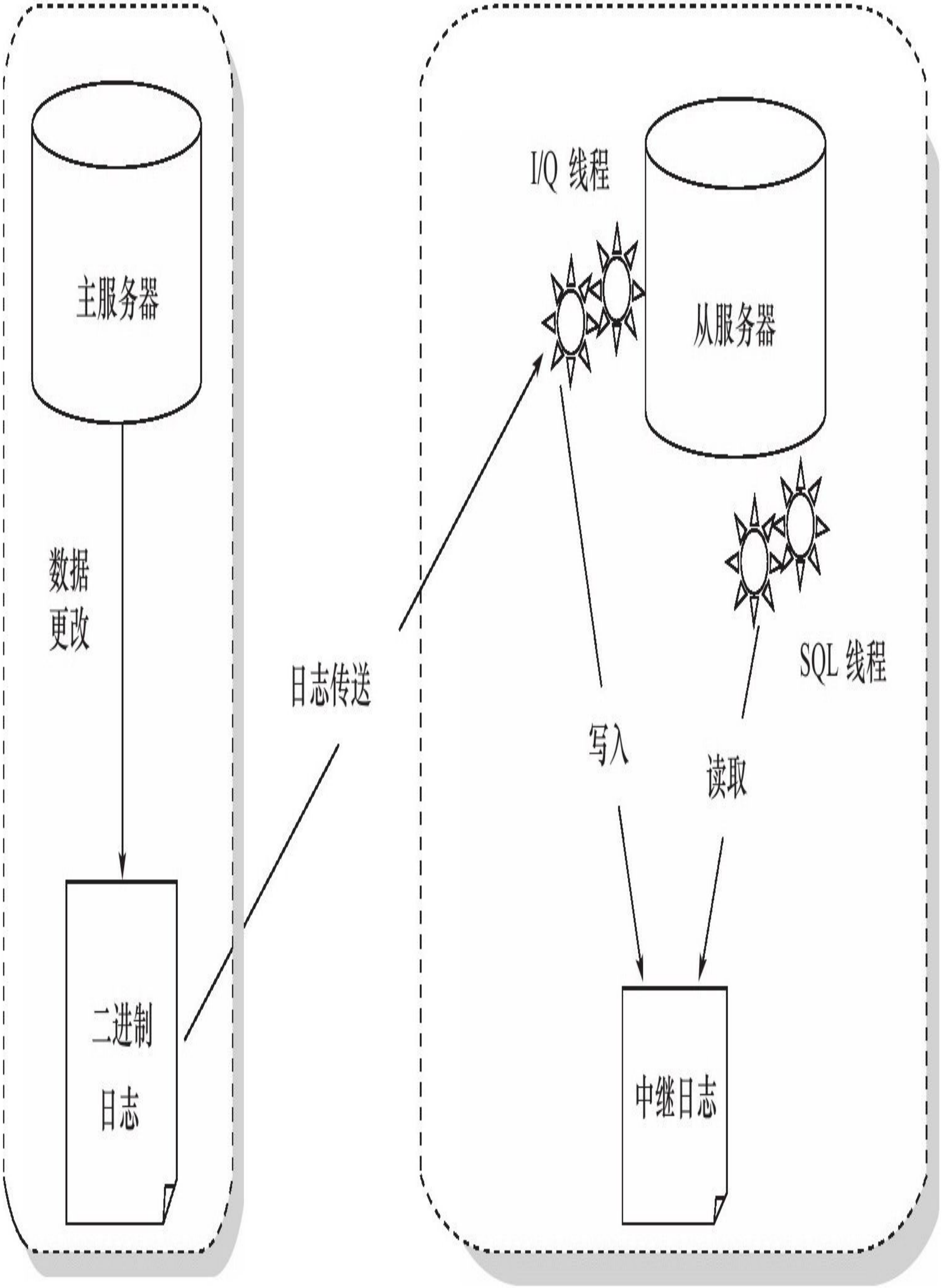
[illegible]replication[illegible]MySQL[illegible] [illegible]replication[illegible]3[illegible]

1[illegible]master[illegible]binlog[illegible]

2[illegible]slave[illegible]relay log[illegible]

3[illegible]

[illegible]8-4[illegible]


主服务器
数据更改
二进制日志
日志传送
I/Q 线程
从服务器
SQL 线程
写入
读取
中继日志

[illegible] 8-4 MySQL[illegible]

[illegible]2[illegible]I/O[illegible]SQL[illegible]MySQL4.0[illegible]1[illegible]SQL[illegible]

```
mysql[illegible]SHOW FULL PROCESSLIST\G;
***************************1.row***************************
Id:1
User:system user
Host:
db:NULL
Command:Connect
Time:6501
State:Waiting for master to send event
Info:NULL
***************************2.row***************************
Id:2
User:system user
Host:
db:NULL
Command:Connect
Time:0
State:Has read all relay log;waiting for the slave I/O thread to update it
Info:NULL
***************************3.row***************************
Id:206
User:root
Host:localhost
db:NULL
```

```
Command:Query

Time:0

State:NULL

Info:SHOW FULL PROCESSLIST

3 rows in set(0.00 sec)
```

□□□□ID□1□□□□□I/O□□□□□□□□□□□□□□□□□□□□□□□ID□2□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□I/O□□□□□

□replication□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
mysql□SHOW FULL PROCESSLIST\G;

……

***************************65.row***************************

Id:26541

User:rep

Host:192.168.190.98:39549

db:NULL

Command:Binlog Dump

Time:6857

State:Has sent all binlog to slave;waiting for binlog to be updated

Info:NULL

……
```

□□□□□□MySQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□SHOW SLAVE STATUS□SHOW MASTER STATUS□□□□□

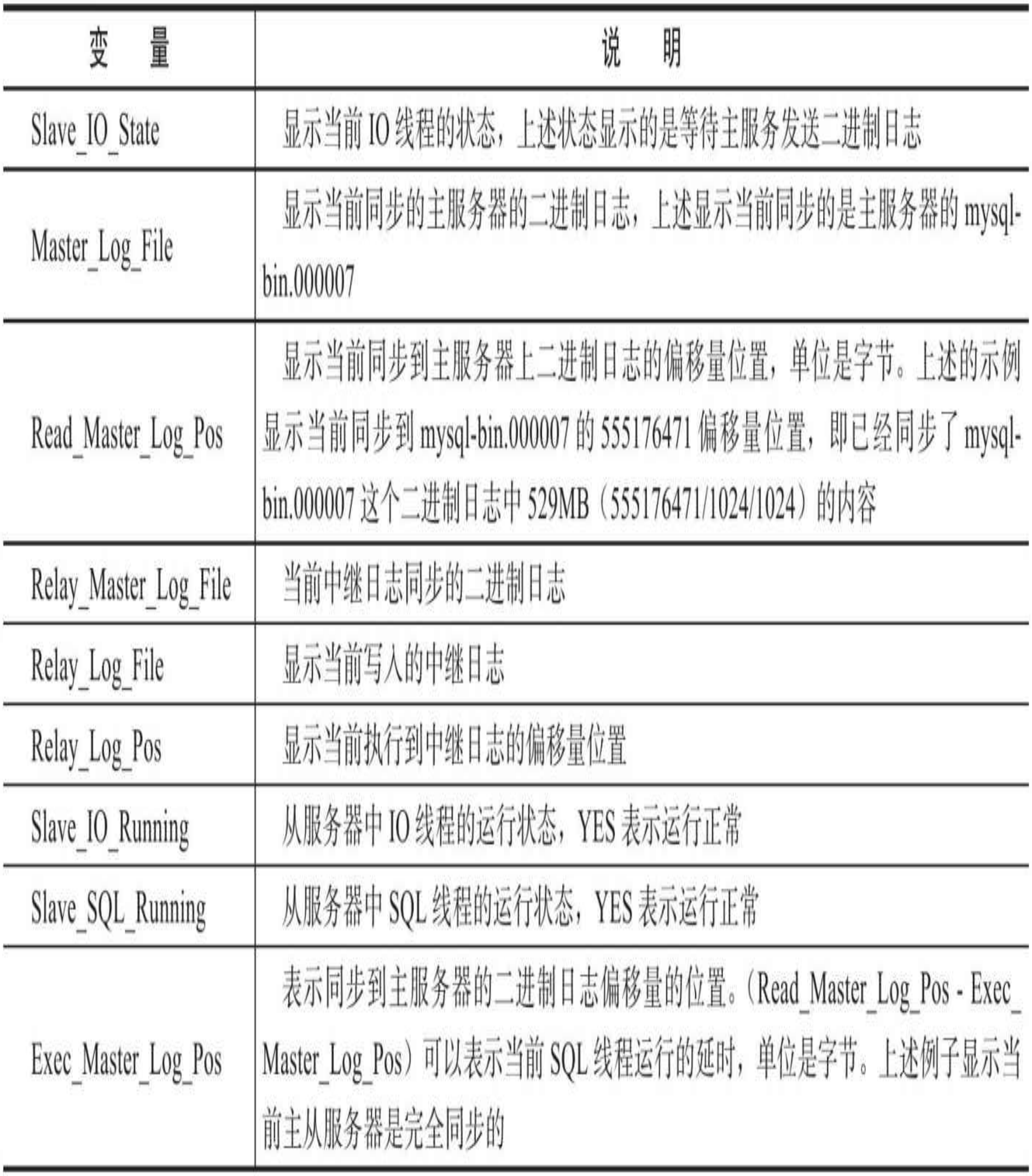
```
mysql□SHOW SLAVE STATUS\G;

***************************1.row***************************

Slave_IO_State:Waiting for master to send event

Master_Host:192.168.190.10
```

```
Master_User:rep
Master_Port:3306
Connect_Retry:60
Master_Log_File:mysql-bin.000007
Read_Master_Log_Pos:555176471
Relay_Log_File:gamedb-relay-bin.000048
Relay_Log_Pos:224355889
Relay_Master_Log_File:mysql-bin.000007
Slave_IO_Running:Yes
Slave_SQL_Running:Yes
Replicate_Do_DB:
Replicate_Ignore_DB:
Replicate_Do_Table:
Replicate_Ignore_Table:
Replicate_Wild_Do_Table:
Replicate_Wild_Ignore_Table:mysql.%,DBA.%
Last_Errno:0
Last_Error:
Skip_Counter:0
Exec_Master_Log_Pos:555176471
Relay_Log_Space:224356045
Until_Condition:None
Until_Log_File:
Until_Log_Pos:0
Master_SSL_Allowed:No
Master_SSL_CA_File:
Master_SSL_CA_Path:
Master_SSL_Cert:
Master_SSL_Cipher:
Master_SSL_Key:
```

```
Seconds_Behind_Master:0

Master_SSL_Verify_Server_Cert:No

Last_IO_Errno:0

Last_IO_Error:

Last_SQL_Errno:0

Last_SQL_Error:

1 row in set(0.00 sec)
```

□□SHOW SLAVE STATUS□□□□□□□□□□□□□□□□□□□□□□□□□8-1□□□

表 8-1　SHOW SLAVE STATUS 的主要变量



变　量	说　明
Slave_IO_State	显示当前 IO 线程的状态，上述状态显示的是等待主服务发送二进制日志
Master_Log_File	显示当前同步的主服务器的二进制日志，上述显示当前同步的是主服务器的 mysql-bin.000007
Read_Master_Log_Pos	显示当前同步到主服务器上二进制日志的偏移量位置，单位是字节。上述的示例显示当前同步到 mysql-bin.000007 的 555176471 偏移量位置，即已经同步了 mysql-bin.000007 这个二进制日志中 529MB（555176471/1024/1024）的内容
Relay_Master_Log_File	当前中继日志同步的二进制日志
Relay_Log_File	显示当前写入的中继日志
Relay_Log_Pos	显示当前执行到中继日志的偏移量位置
Slave_IO_Running	从服务器中 IO 线程的运行状态，YES 表示运行正常
Slave_SQL_Running	从服务器中 SQL 线程的运行状态，YES 表示运行正常
Exec_Master_Log_Pos	表示同步到主服务器的二进制日志偏移量的位置。(Read_Master_Log_Pos - Exec_Master_Log_Pos) 可以表示当前 SQL 线程运行的延时，单位是字节。上述例子显示当前主从服务器是完全同步的

□□SHOW MASTER STATUS□□□□□□□□□□□□□□□□□□□□□□

```
mysql□SHOW MASTER STATUS\G;
```

```
***************************1.row***************************

File:mysql-bin.000007

Position:606181078

Binlog_Do_DB:

Binlog_Ignore_DB:

1 row in set(0.01 sec)
```

[illegible]606181078[illegible] Read_Master_Log_Pos[illegible]I/O[illegible]

[illegible]MySQL[illegible]I/O[illegible]SQL[illegible]

8.7.2 [illegible]+[illegible]

[illegible]

❑[illegible]MySQL[illegible]

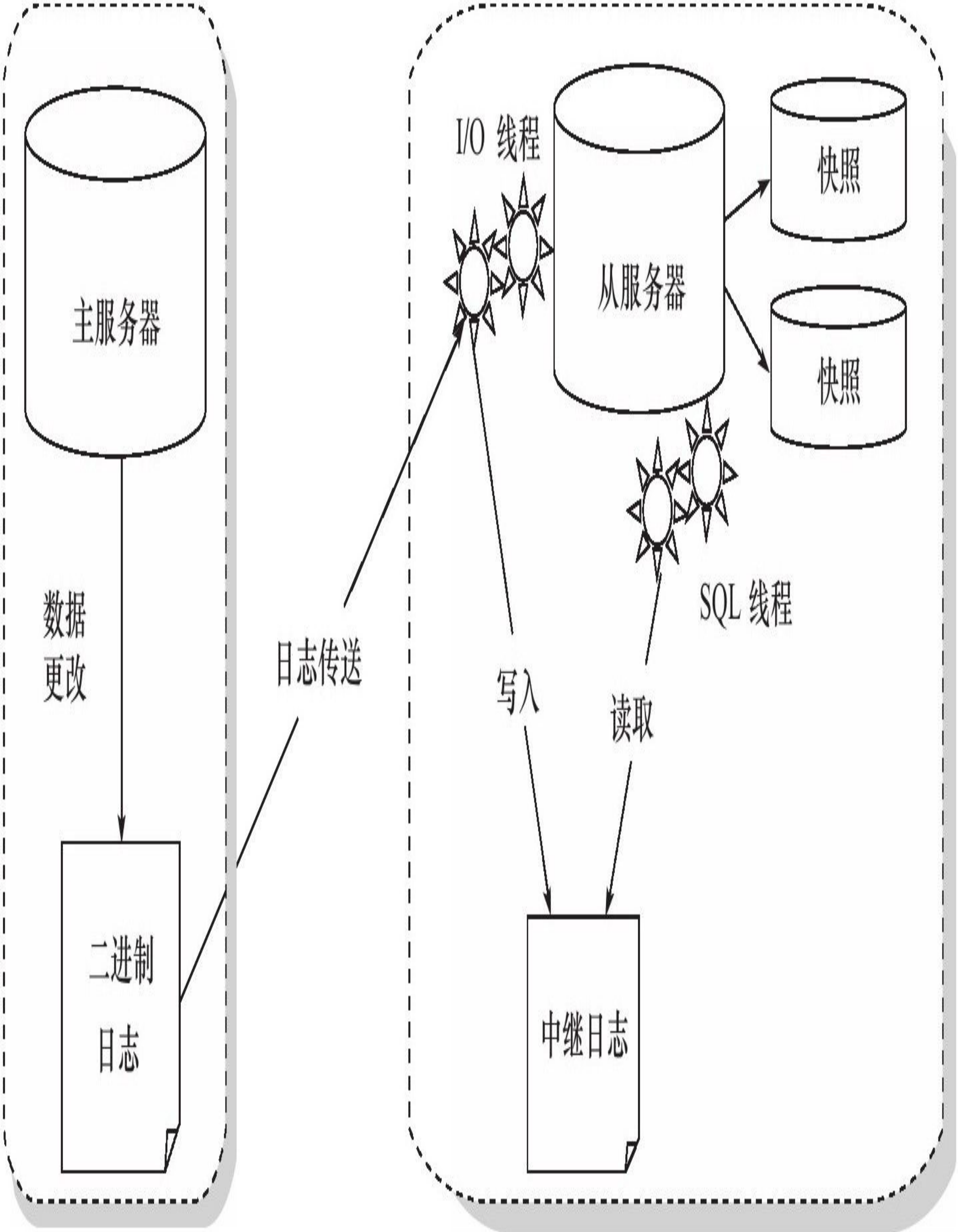
❑[illegible]DNS[illegible]Round-Robin[illegible]Linux[illegible]LVS[illegible]

❑[illegible]

❑[illegible]

[illegible]DBA[illegible]DROP DATABASE[illegible] DROP TABLE[illegible]

[illegible]point-in-time [illegible]+[illegible]8-5[illegible]



[illegible] 8-5 [illegible]+[illegible]

[illegible]

[illegible]read-only[illegible]

```
[mysqld]

read-only
```

[illegible]read-only[illegible]SUPER[illegible]

```
mysql[illegible]INSERT INTO z SELECT 2;

ERROR 1290(HY000):The MySQL server is running with the--read-only option so it cannot execute this statement
```

8.8 [illegible]

[illegible]MySQL[illegible]InnoDB[illegible]mysqldump[illegible]xtrabackup[illegible]InnoDB[illegible]MySQL[illegible]

[illegible]9[illegible] [illegible]

[illegible]InnoDB[illegible]InnoDB[illegible]InnoDB[illegible]

❑[illegible]CPU

❑[illegible]

❑[illegible]

❑[illegible]RAID

❑[illegible]

❑[illegible]

❑[illegible]

9.1 [illegible]CPU

[illegible]OLTP[illegible]Online Transaction Processing[illegible]OLAP[illegible]Online Analytical Processing[illegible]OLAP[illegible]SQL[illegible]OLTP[illegible]Blog[illegible]OLAP[illegible]

InnoDB[illegible]OLTP[illegible]

❑[illegible]

❑[illegible]

❑[illegible]

❑[illegible]

[illegible]OLTP[illegible]CPU[illegible]CPU[illegible]OLTP[illegible]OLAP[illegible]CPU[illegible]OLTP[illegible]IO[illegible]IO[illegible]

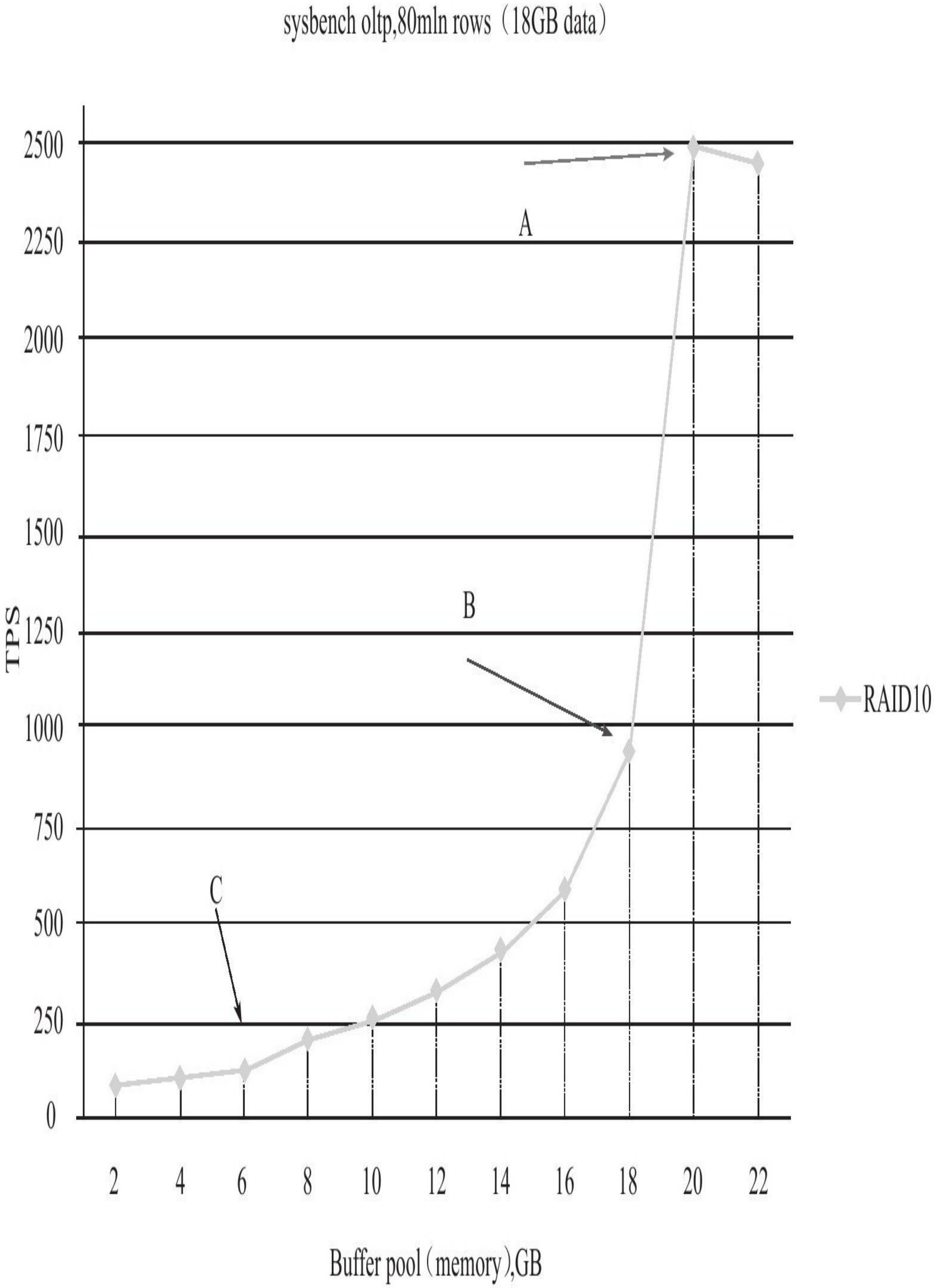
[illegible]CPU[illegible]64[illegible]64[illegible]64[illegible]CPU[illegible]4[illegible]CPU[illegible]Intel[illegible]AMD[illegible]8[illegible]CPU[illegible]128[illegible]CPU[illegible]

[illegible]InnoDB[illegible]master thread[illegible]InnoDB1.0[illegible]CPU[illegible]InnoDB 1.2[illegible]purge[illegible]master thread[illegible]CPU[illegible]InnoDB[illegible]1.1[illegible]CPU[illegible]innodb_read_io_threads[illegible]innodb_write_io_threads[illegible]IO[illegible]CPU[illegible]

□□□□MySQL□□□□□□□□□SQL□□□□□□□□□CPU□□□□□□□□□CPU□□□□
OLTP□□□□□□□□□□□□□□□□□□OLTP□□□□□□□□□□□□□□□□CPU□□□
CPU□□□□□□□□□□□□□□□□□

9.2 [illegible]

[illegible]InnoDB[illegible]InnoDB Buffer Pool[illegible]Percona[illegible]CTO Vadim[illegible]9-1[illegible]


sysbench oltp,80mln rows（18GB data）
TPS
2500
2250
2000
1750
1500
1250
1000
750
500
250
0
A
B
C
RAID10
2
4
6
8
10
12
14
16
18
20
22
Buffer pool（memory),GB

□ 9-1 □□□□□□□InnoDB□□□□□□□□□

□□□□□□□□□□□□□□□□18GB□□□□□□□□□□□□□□2GB□4GB□6GB□8GB□10GB□12GB□14GB□16GB□18GB□20GB□22GB□□□□sysbench□□□□□□□□□□□□□□□□□□□□TPS□Transaction Per Second□□□□□□□□□□□□□□20GB□22GB□□

□□□□□□□□□□□□□“□□”□□□64□□□□□□□

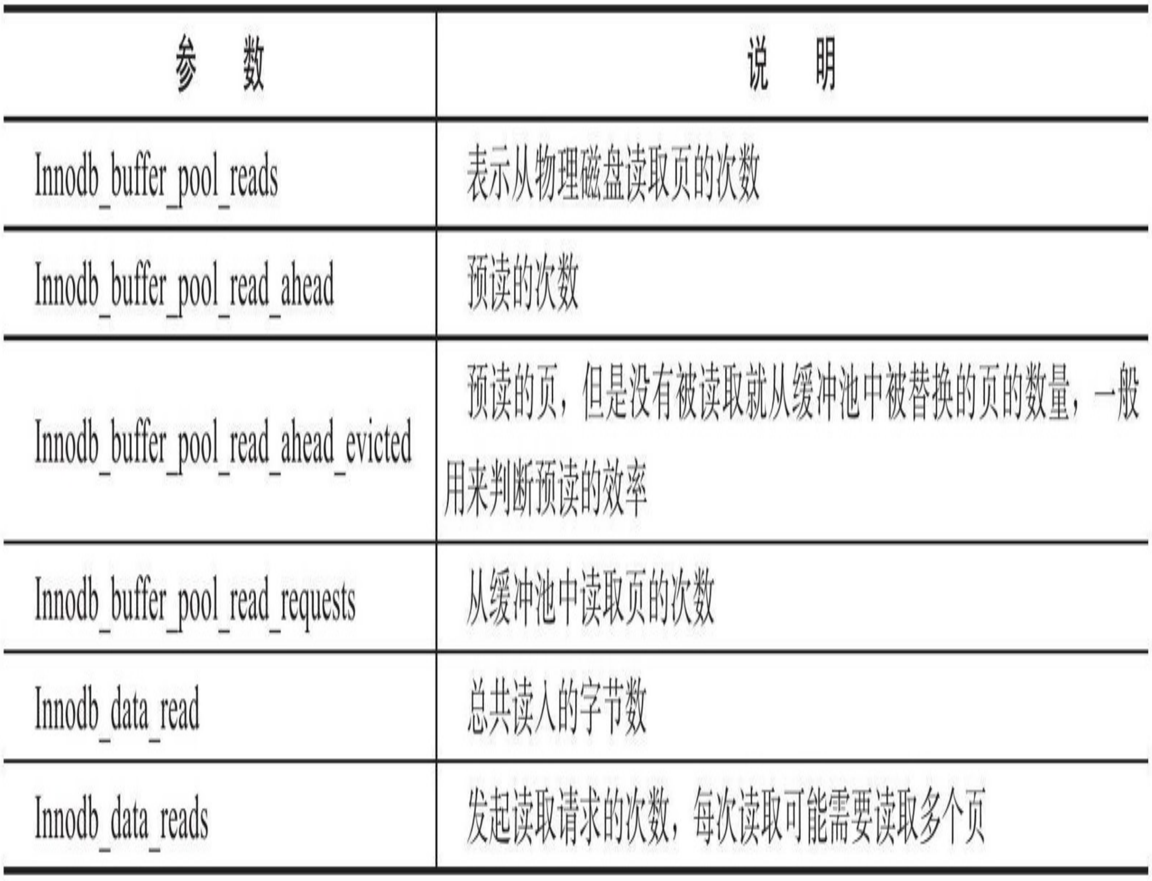
□□InnoDB□□□□□□□□□□□□□□□□□99%□□□

```
mysql□SHOW GLOBAL STAUTS LIKE'innodb%read%'\G;

***************************1.row***************************

Variable_name:Innodb_buffer_pool_read_ahead

Value:0

***************************2.row***************************

Variable_name:Innodb_buffer_pool_read_ahead_evicted

Value:0

***************************3.row***************************

Variable_name:Innodb_buffer_pool_read_requests

Value:167051313

***************************4.row***************************

Variable_name:Innodb_buffer_pool_reads

Value:129236

***************************5.row***************************

Variable_name:Innodb_data_pending_reads

Value:0

***************************6.row***************************

Variable_name:Innodb_data_read

Value:2135642112
```

```
***************************7.row***************************

Variable_name:Innodb_data_reads

Value:130309

***************************8.row***************************

Variable_name:Innodb_pages_read

Value:130215

***************************9.row***************************

Variable_name:Innodb_rows_read

Value:17651085

9 rows in set(0.00 sec)
```

□□□□□□□□□□□9-1□□□

表 9-1 当前服务器的状态参数



参　　数	说　　明
Innodb_buffer_pool_reads	表示从物理磁盘读取页的次数
Innodb_buffer_pool_read_ahead	预读的次数
Innodb_buffer_pool_read_ahead_evicted	预读的页，但是没有被读取就从缓冲池中被替换的页的数量，一般用来判断预读的效率
Innodb_buffer_pool_read_requests	从缓冲池中读取页的次数
Innodb_data_read	总共读入的字节数
Innodb_data_reads	发起读取请求的次数，每次读取可能需要读取多个页

□□□□□□□□□□□□□□□□□□

$$\text{缓冲池命中率} = \frac{\text{Innodb\_buffer\_pool\_read\_requests}}{\text{（Innodb\_buffer\_pool\_read\_requests+Innodb\_buffer\_pool\_read\_ahead+Innodb\_buffer\_pool\_reads）}}$$

$$\text{平均每次读取的字节数} = \frac{\text{Innodb\_data\_read}}{\text{Innodb\_data\_reads}}$$

[illegible]=167 051 313/[illegible]167 051 313+129 236+0[illegible]
=99.92%[illegible]

[illegible]

9.3 [illegible]

9.3.1 [illegible]

[illegible]SAS[illegible]SATA[illegible]2.5[illegible]SAS[illegible]

[illegible]3ms[illegible]15 000RPM[illegible]rotate per minute[illegible]

[illegible]RAID[illegible]

9.3.2 [illegible]

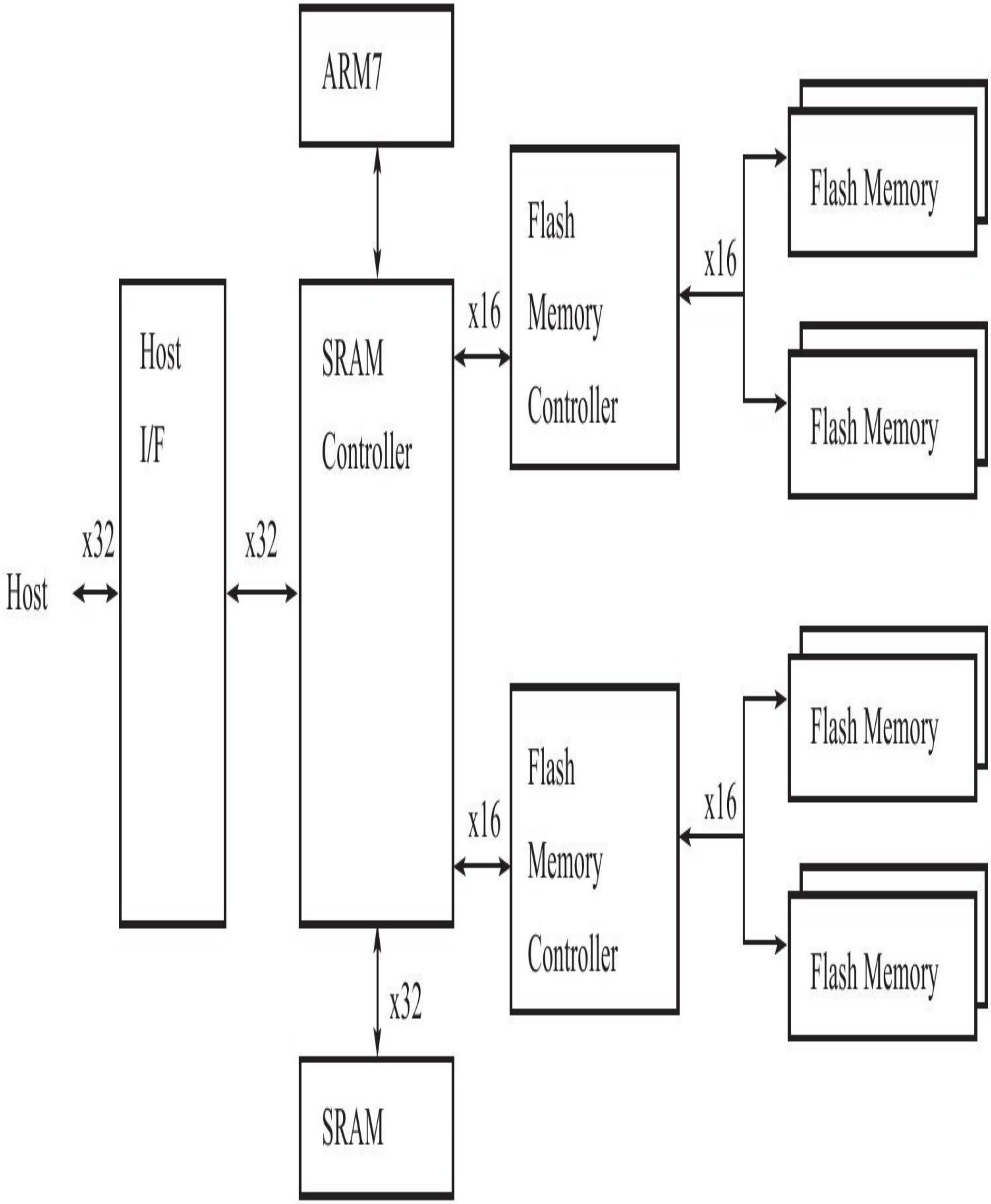
[illegible]Flash Memory[illegible]EMC[illegible]TB[illegible]Oracle[illegible]Exadata[illegible]

[illegible]

[illegible]sector[illegible]erase[illegible]128KB[illegible]256KB[illegible]

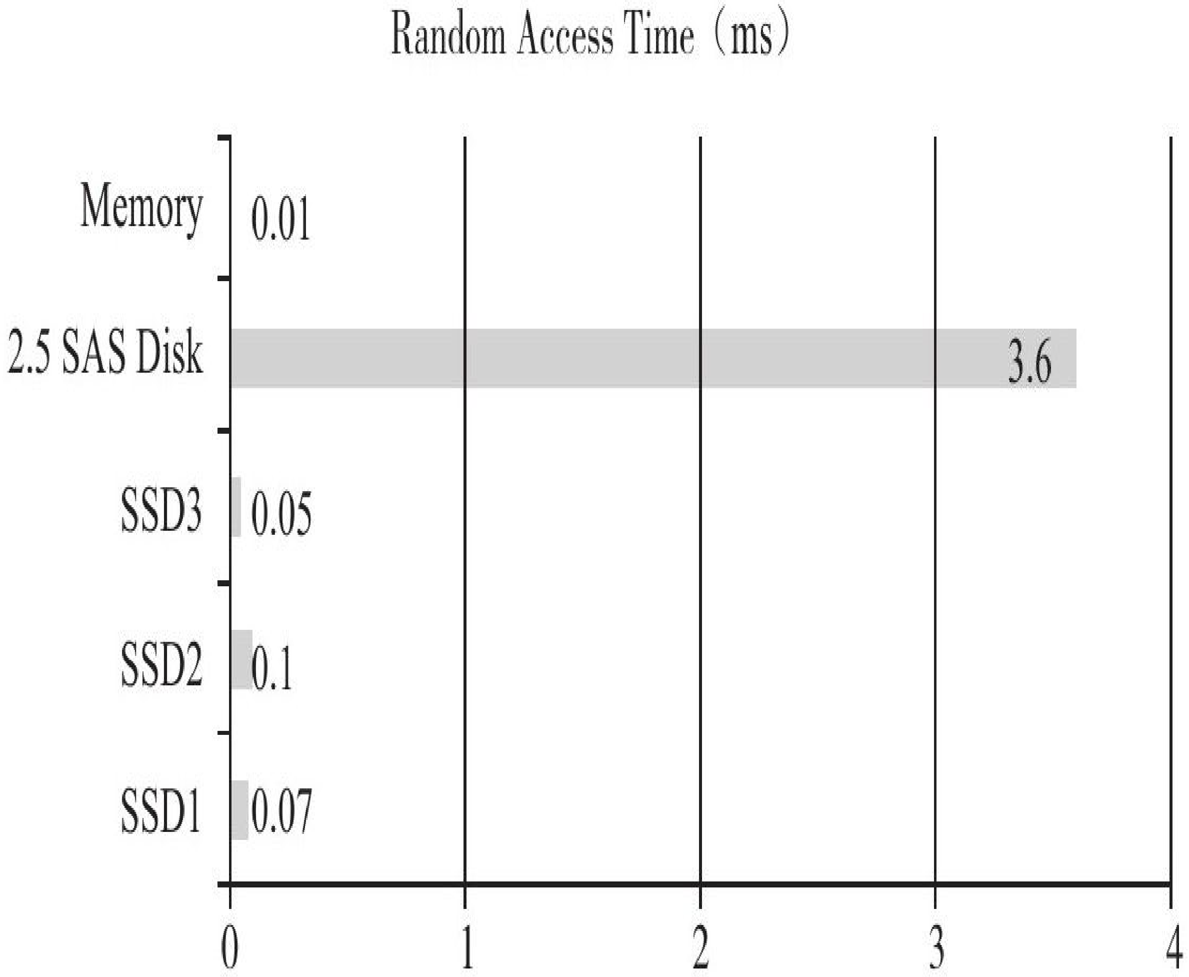
[illegible]

[illegible]9-2[illegible]4[illegible]Intel X-25M[illegible]10[illegible]


ARM7
Host
x32
Host
I/F
x32
SRAM
Controller
x16
Flash
Memory
Controller
x16
Flash Memory
Flash Memory
x16
Flash
Memory
Controller
x16
Flash Memory
Flash Memory
x32
SRAM

[illegible] 9-2 [illegible]

[illegible]I/O[illegible]0.1ms[illegible]9-3[illegible]



[illegible] 9-3 [illegible]

[illegible]InnoDB[illegible]innodb_io_capacity[illegible]IOPS[illegible]InnoSQL[illegible]InnoDB1.2[illegible]

[illegible]InnoSQL[illegible]L2 Cache[illegible]

[illegible]Cache[illegible]Facebook Flash Cache[illegible]bcache[illegible]InnoDB[illegible]

9.4 [illegible]RAID

9.4.1 RAID[illegible]

RAID[illegible]Redundant Array of Independent Disks[illegible]RAID[illegible]

RAID[illegible]

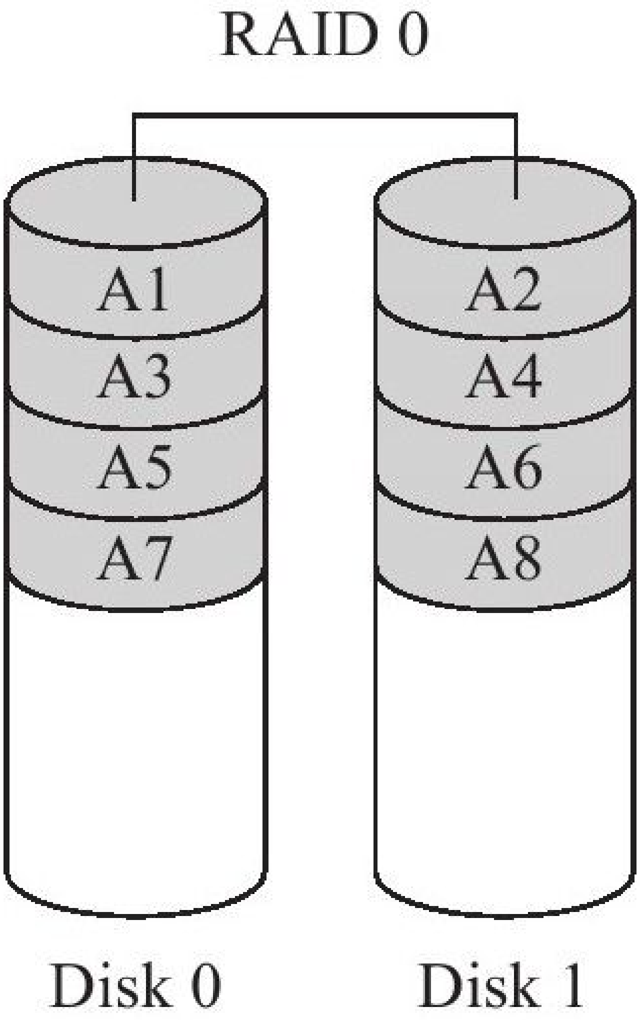
❑[illegible]

❑[illegible]

❑[illegible]

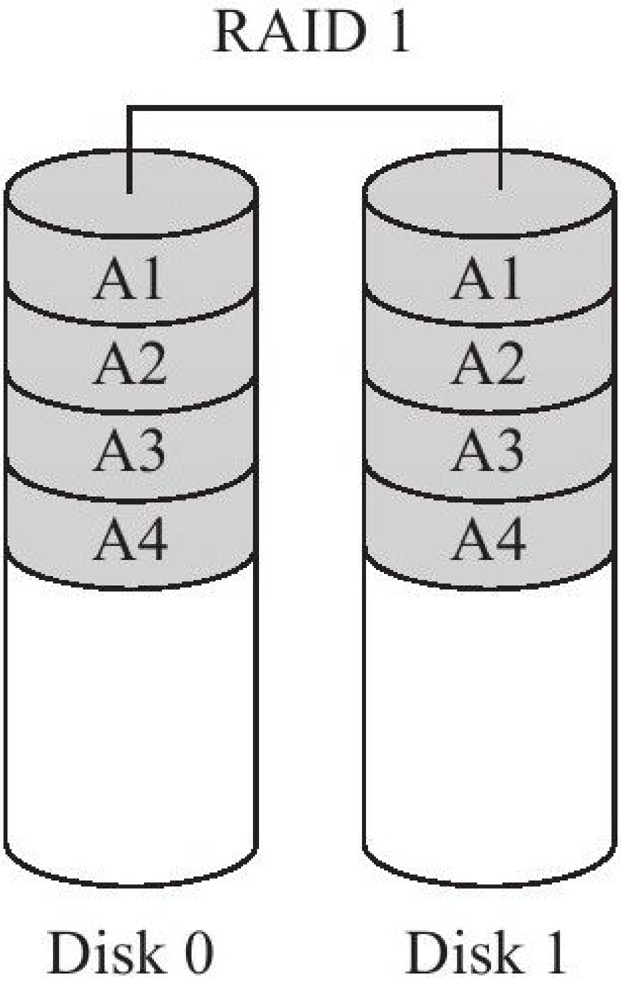
[illegible]RAID[illegible]RAID 0[illegible]RAID 1[illegible]RAID 5[illegible]RAID 10[illegible]RAID 50[illegible]

RAID 0[illegible]I/O[illegible]RAID 0[illegible]9-4[illegible]RAID 0[illegible]RAID 0[illegible]×[illegible]I/O[illegible]RAID[illegible]50MB/s[illegible]RAID 0[illegible]96MB/s[illegible]RAID 0[illegible]130MB/s[illegible]150MB/s[illegible]



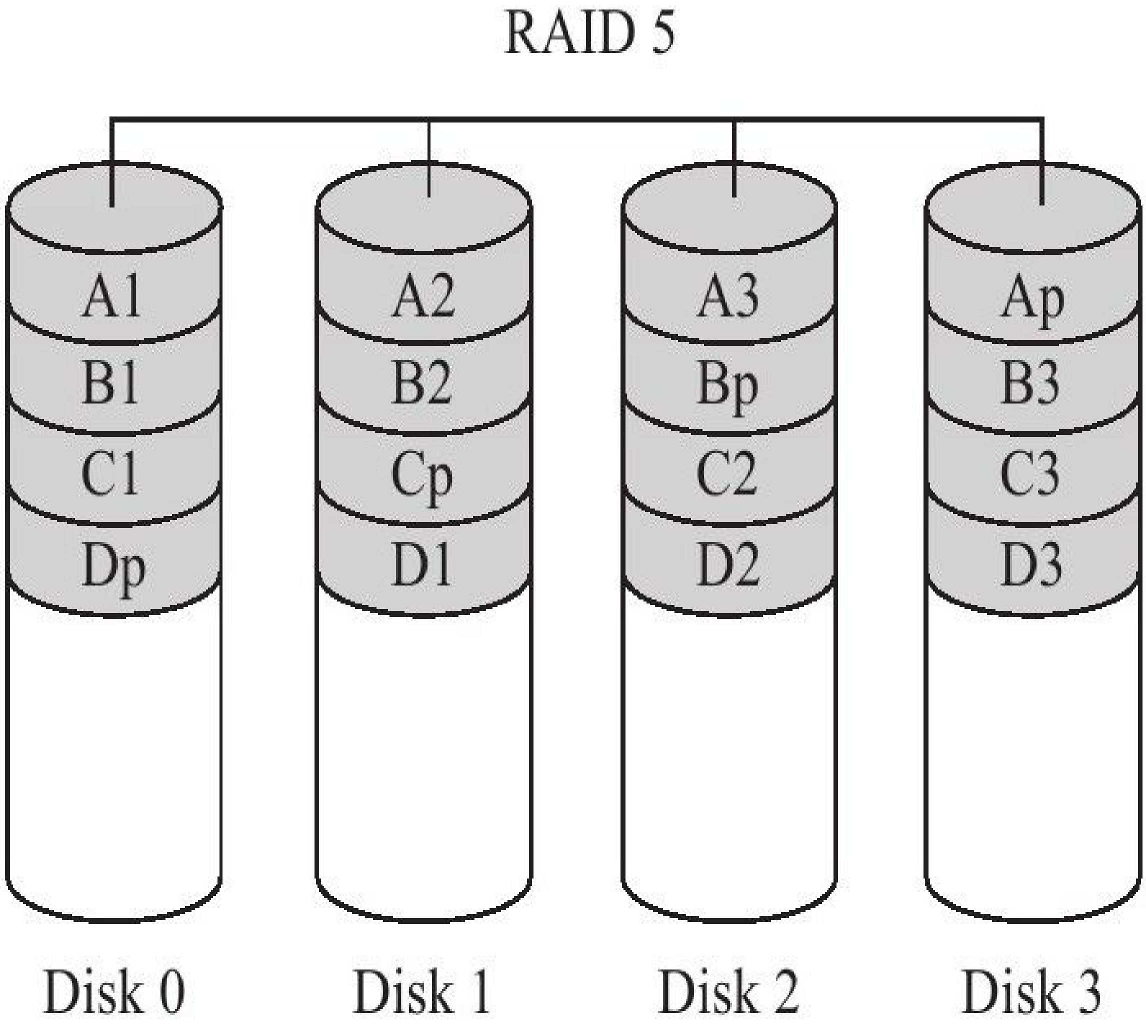
[illegible] 9-4 RAID 0[illegible]

RAID 1[illegible]N[illegible]9-5[illegible]RAID 1[illegible]RAID 1[illegible]RAID[illegible]RAID 1[illegible]RAID[illegible]



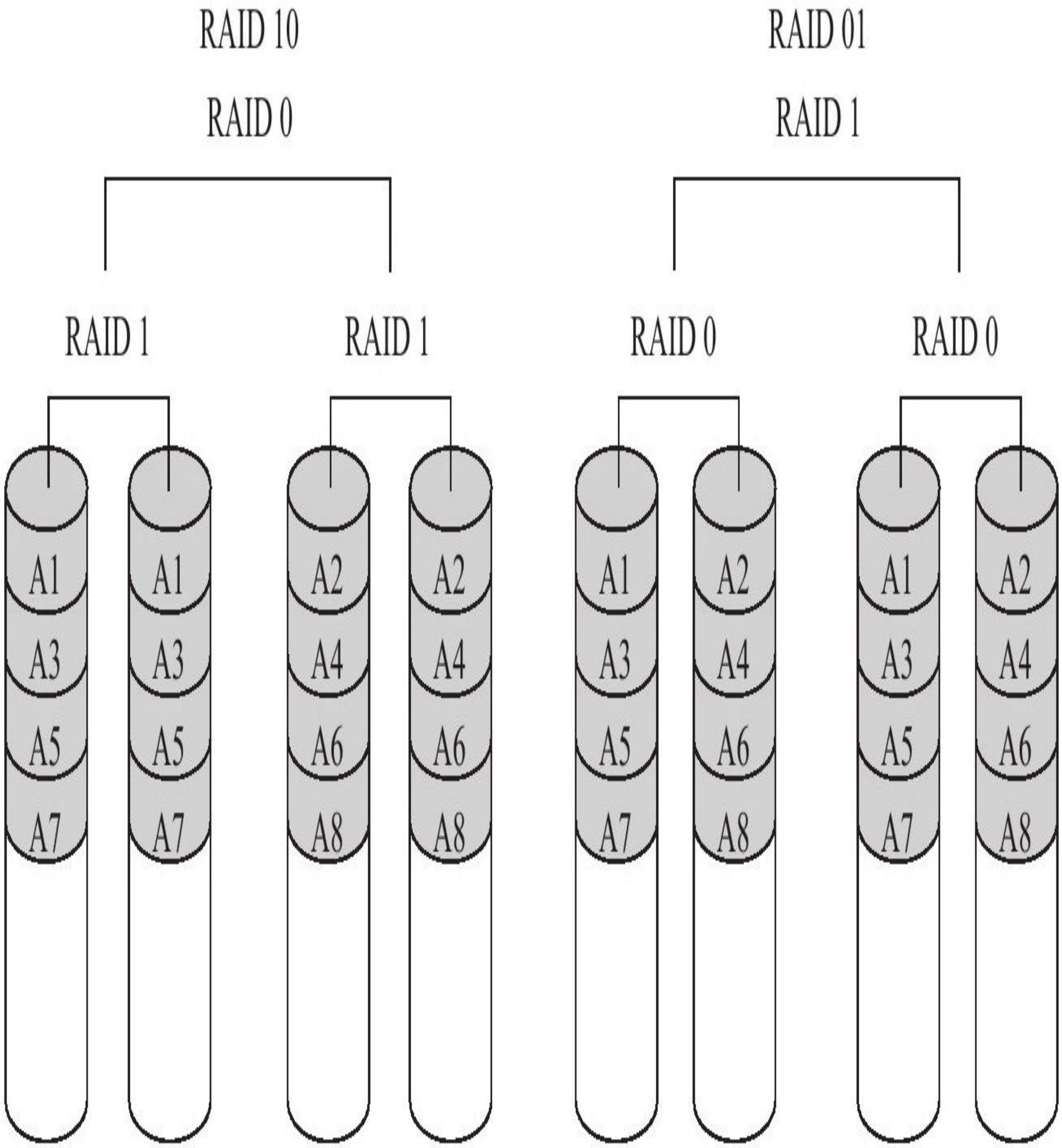
[illegible] 9-5 RAID 1[illegible]

RAID 5[illegible]Disk Striping[illegible]RAID 5[illegible]RAID 5[illegible]RAID 5[illegible]RAID 5[illegible]RAID 5[illegible]RAID 0[illegible]RAID 1[illegible]RAID 5[illegible]RAID 5[illegible]RAID 0[illegible]Write Back[illegible]RAID 5[illegible]RAID 1[illegible]RAID 5[illegible]9-6[illegible]



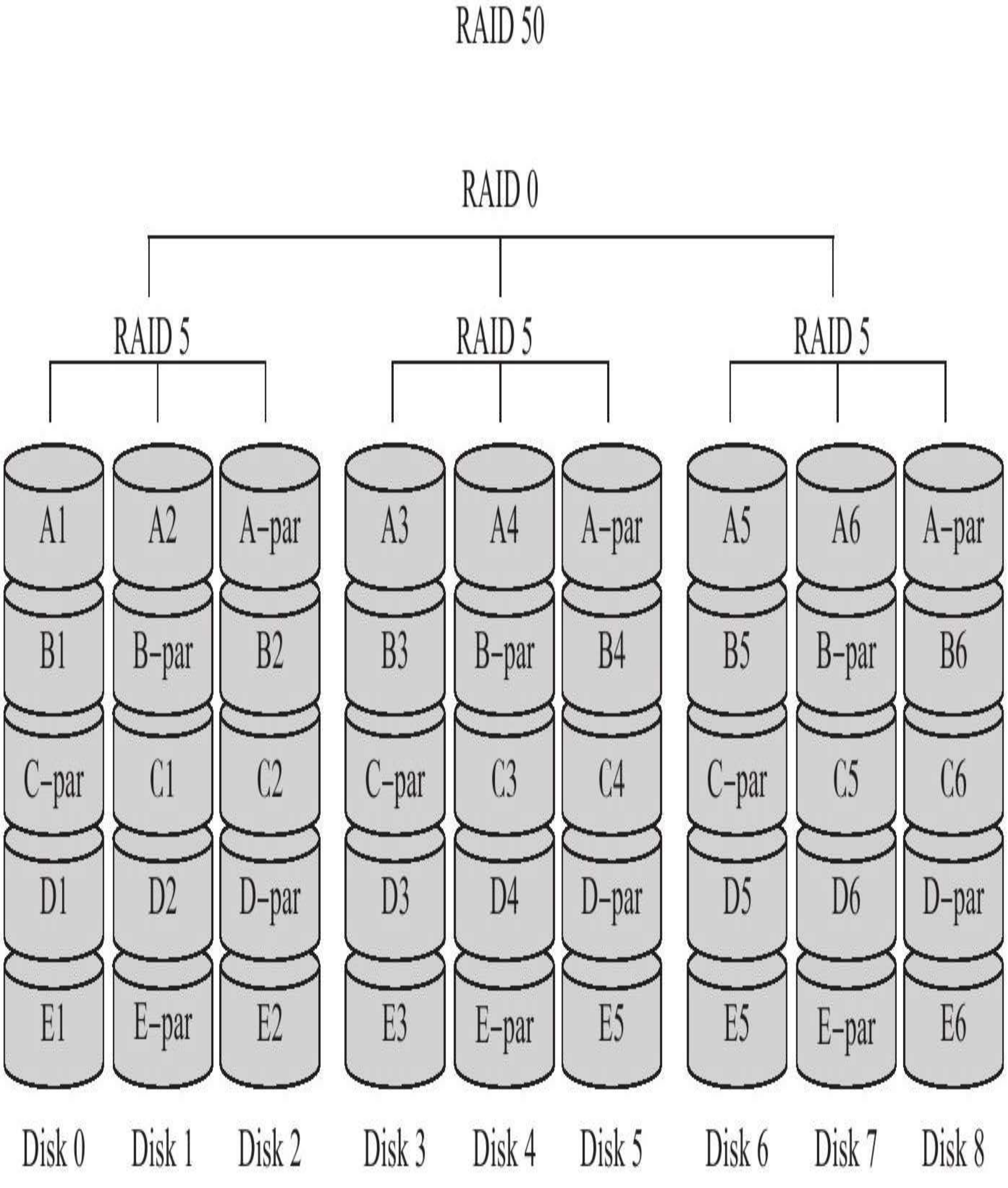
□ 9-6 RAID 5□□

RAID 10□RAID 01□RAID 10□□□□□□□□□□□□□□□□□□□□□RAID 0□□□□□□□□□□□□□□□RAID 1□□□RAID 10□□□□□□□□□□□□□□RAID 0□□□□□□□□□RAID 01□□RAID 10□□□□□□□□□□□□□□□□□□□□RAID 01□□□□□□□□□□□□RAID 1□□□□□□□□□□□□□□□RAID 0□□□RAID 01□RAID 10□□□RAID 01□□□□□□□□RAID 10□□□□□RAID 0□□□□□□□□□RAID 10□□□□□□RAID 0□□□□RAID 1□□□RAID 10□RAID 01□□□□□9-7□□□



□ 9-7 RAID 10□RAID 01□□

RAID 50□RAID 50□□□□□□□□□□□□□□□□□□□□□RAID 0□□□□□□□□□□□□□□□□□□□□□□□□□RAID 5□□RAID 5□□□□□□



[illegible] 9-8 RAID 50[illegible]

[illegible]RAID 10[illegible]RAID 1[illegible]RAID 0[illegible]strip[illegible]

□□□□□□□□□□□□□□□□□□□□□□□□□IO□□□20%□□□□□□□□□□IO□□□□□
90%□

9.4.2 RAID Write Back[illegible]

RAID Write Back[illegible]RAID[illegible]sync_binlog[illegible]1[illegible]

[illegible]Write Back[illegible]RAID[illegible]RAID[illegible]BBU[illegible]Battery Backup Unit[illegible]Write Back[illegible]RAID[illegible]

[illegible]Write Back[illegible]RAID[illegible]Write Through[illegible]Write Through[illegible]

[illegible]Write Back[illegible]RAID[illegible]Write Through[illegible]Write Back[illegible]RAID[illegible]RAID[illegible]Write Back[illegible]Write Through[illegible]

[illegible]Write Back[illegible]Write Back[illegible]Write Back[illegible]

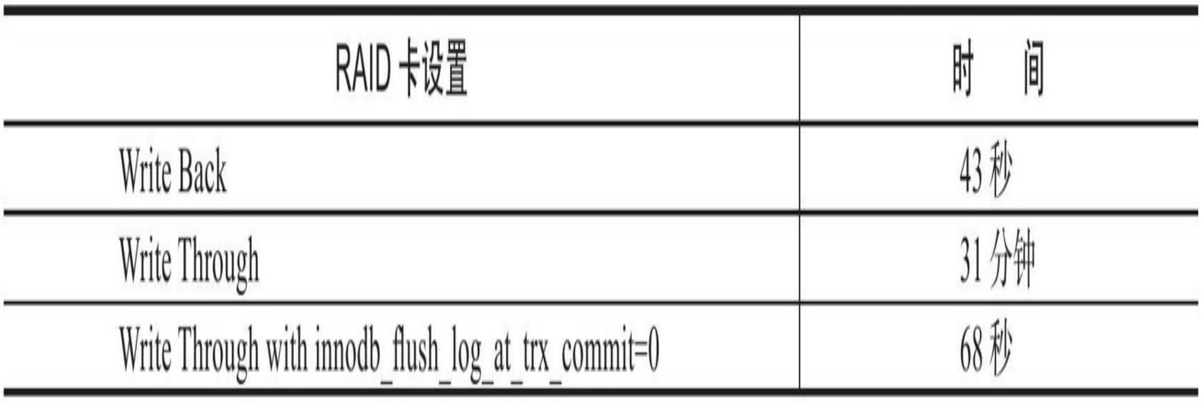
[illegible]20W[illegible]Write Back[illegible]Write Through[illegible]

```
mysql□CREATE TABLE t(a CHAR(2))Engine=InnoDB;

Query OK,0 rows affected(0.00 sec)

mysql□DELIMITER//

mysql□

mysql□CREATE PROCEDURE p()

-□BEGIN

-□DECLARE v INT;
```

```
-□SET v=0;

-□WHILE v□200000 DO

-□INSERTINTO t VALUES('aa');

-□SET v=v+1;

-□END WHILE;

-□END

-□//

Query OK,0 rows affected(0.12 sec)

mysql□DELIMITER;
```

□□□□□□□□t□□20W□□□□□□□□□□Write Back□Write Through□□□□□□□□□□□□□□□□□□□9-2□□□

表 9-2 Write Back 和 Write Through 的性能对比测试结果



RAID 卡设置	时　间
Write Back	43 秒
Write Through	31 分钟
Write Through with innodb_flush_log_at_trx_commit=0	68 秒

□□□□□□□□□□□□□□□□□□□□□□□□□CALL P□□□□□□□□□□□□□□□20W□□□□□□□□□□□□□□Write Back□□□□□□□□□□43□□□□Write Through□□□□□□□□□31□□□□□□40□□□□□□

□□□□Write Through□□□□□□□□□innodb_flush_log_at_trx_commit□□□0□□□□□□□□□□□P□□□□□□□□□68□□master□□□

9.4.3 RAID□□□□

□RAID□□□□□□□□□□□□□□□□□□□□□BIOS□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□RAID□□□□□□□□□□□□□LSI□□□□□□□RAID□□□□□□□MegaCLI□□□□□□□□

MegaCLI□□□□□□□□□□□□□□Windows□□□□□□□□GUI□□□□□□□□□□□□□□□□□□□□□□□□□□□□□MegaCLI□□□□□Windows□□□□□□□□□MegaCLI.exe□

□□MegaCLI□□RAID□□□□□

```
[root@xen-server□]#/opt/MegaRAID/MegaCli/MegaCli64-AdpAllInfo-a0

Adapter#0

==============================================================================

Versions

================

Product Name:MegaRAID SAS 8708ELP

Serial No:P012233608

FW Package Build:9.0.1-0030

……

HW Configuration

================

SAS Address:500605b000d1e180

BBU:Present

Alarm:Present

NVRAM:Present

Serial Debugger:Present

Memory:Present

Flash:Present

Memory Size:256MB
```

```
TPM:Absent

Default Settings

================

Phy Polarity:0

PhyPolaritySplit:240

Background Rate:30

Stripe Size:64kB

Flush Time:4 seconds

Write Policy:WB

Read Policy:None

Cache When BBU Bad:Disabled

Cached IO:No

SMART Mode:Mode 6

Alarm Disable:Yes

Coercion Mode:1GB

ZCR Config:Unknown

Dirty LED Shows Drive Activity:No

BIOS Continue on Error:No

Spin Down Mode:None

Allowed Device Type:SAS/SATA Mix

Allow Mix in Enclosure:Yes

Allow HDD SAS/SATA Mix in VD:Yes

Allow SSD SAS/SATA Mix in VD:No

Allow HDD/SSD Mix in VD:No

Allow SATA in Cluster:No

Max Chained Enclosures:3

Disable Ctrl-R:Yes

Enable Web BIOS:Yes

Direct PD Mapping:No

BIOS Enumerate VDs:Yes
```

```
Restore Hot Spare on Insertion:No

Expose Enclosure Devices:Yes

Maintain PD Fail History:Yes

Disable Puncturing:No

Zero Based Enclosure Enumeration:No

PreBoot CLI Enabled:Yes

LED Show Drive Activity:No

Cluster Disable:Yes

SAS Disable:No

Auto Detect BackPlane Enable:SGPIO/i2c SEP

Use FDE Only:No

Enable Led Header:No

Delay during POST:0
```

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□RAID□□□□□□□□□□□□RAID□□□□□MegaRAID SAS 8708ELP□□□□□□256MB□□□□□□□□□□□□□□□□□□□□Write Policy□WB□Write Back□□□

MegaCLI□□□□□□□□□□□□□□□□□□□

```
[root@xen-server□]#/opt/MegaRAID/MegaCli/MegaCli64-PDList-aALL

Adapter#0

Enclosure Device ID:252

Slot Number:0

Device Id:8

Sequence Number:2

Media Error Count:0

Other Error Count:0

Predictive Failure Count:0

Last Predictive Failure Event Seq Number:0

PD Type:SAS
```

```
Raw Size:279.396 GB[0x22ecb25c Sectors]
Non Coerced Size:278.896 GB[0x22dcb25c Sectors]
Coerced Size:278.464 GB[0x22cee000 Sectors]
Firmware state:Online
SAS Address(0):0x5000c5000f363b55
SAS Address(1):0x0
Connected Port Number:0(path0)
Inquiry Data:SEAGATE ST3300655SS 00023LM5MGZZ
FDE Capable:Not Capable
FDE Enable:Disable
Secured:Unsecured
Locked:Unlocked
Foreign State:None
Device Speed:Unknown
Link Speed:Unknown
Media Type:Hard Disk Device
……
```

[illegible]SEAGATE ST3300655SS[illegible]http://discountechnology.com/Seagate-ST3300655SS-SAS-Hard-Drive[illegible]3.5[illegible]15 000[illegible]Cache[illegible]16MB[illegible]3.5[illegible]4.0[illegible]

[illegible]Write Back[illegible]

```
[root@xen-server[illegible]]#/opt/MegaRAID/MegaCli/MegaCli64-LDGetProp-Cache-LALL-aALL
Adapter 0-VD 0(target id:0):Cache Policy:WriteBack,ReadAheadNone,Direct,No Write Cache if bad BBU
Adapter 0-VD 1(target id:1):Cache Policy:WriteBack,ReadAheadNone,Direct,No Write Cache if bad BBU
Exit Code:0x00
```

[illegible]RAID[illegible]Write Back[illegible]BBU[illegible]Write Back[illegible]RAID[illegible]

[illegible]

```
#/opt/MegaRAID/MegaCli/MegaCli64-LDSetPropWB-LALL-aALL

#/opt/MegaRAID/MegaCli/MegaCli64-LDSetPropWT-LALL-aALL
```

[illegible]RAID[illegible]Write Back[illegible]Write Through[illegible]Write Through[illegible]Write Back[illegible]

9.5 [illegible]

Linux[illegible]MySQL[illegible]Linux[illegible]Linux[illegible]

[illegible]Linux[illegible]FreeBSD[illegible]FreeBSD[illegible]MySQL[illegible]FreeBSD[illegible]MySQL[illegible]

Solaris[illegible]SPARC[illegible]X86[illegible]Solaris[illegible]ZFS[illegible]MySQL[illegible]Open Solaris[illegible]

Windows[illegible]MySQL[illegible]Windows[illegible]MySQL[illegible]Linux[illegible]Windows[illegible]Linux[illegible]

4G[illegible]8G[illegible]4G[illegible]64[illegible]64[illegible]64[illegible]64[illegible]32[illegible]MySQL[illegible]64[illegible]64[illegible]

9.6 [illegible]

[illegible]Windows[illegible]NTFS[illegible]Solaris[illegible]ZFS[illegible]Linux[illegible]EXT3[illegible]ReiserFS[illegible]EXT4[illegible]XFS[illegible]

[illegible]XFS[illegible]“[illegible]”[illegible]EXT3[illegible]EXT3[illegible]DBA[illegible]

[illegible]DBA[illegible]ZFS[illegible]LVM[illegible]mount[illegible]

9.7 [illegible]

[illegible]MySQL[illegible]sysbench[illegible]mysql-tpcc[illegible]

9.7.1 sysbench

sysbench[illegible]

❑CPU[illegible]

❑[illegible]IO[illegible]

❑[illegible]

❑[illegible]

❑POSIX[illegible]

❑[illegible]OLTP[illegible]

sysbench[illegible]OLTP[illegible]MySQL[illegible]PostgreSQL[illegible]Oracle[illegible] sysbench[illegible]Linux[illegible]sysbench[illegible]Windows[illegible]Microsoft SQL Server[illegible]

sysbench[illegible]http://sysbench.sourceforge.net[illegible]sysbench[illegible]Linux[illegible]RED HAT[illegible]sysbench[illegible]

sysbench[illegible]

```
[root@xen-server[illegible]]#sysbench

Missing required command argument.

Usage:
```

```
sysbench[general-options]...--test=□test-name□[test-options]...command

General options:

--num-threads=N number of threads to use[1]

--max-requests=N limit for total number of requests[10000]

--max-time=N limit for total execution time in seconds[0]

--thread-stack-size=SIZE size of stack per thread[32K]

--init-rng=[on|off]initialize random number generator[off]

--test=STRING test to run

--debug=[on|off]print more debugging info[off]

--validate=[on|off]perform validation checks where possible[off]

--help=[on|off]print help and exit

--version=[on|off]print version and exit

Compiled-in tests:

fileio-File I/O test

cpu-CPU performance test

memory-Memory functions speed test

threads-Threads subsystem performance test

mutex-Mutex performance test

oltp-OLTP test

Commands:prepare run cleanup help version

See'sysbench--test=□name□help'for a list of options for each test.
```

□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□OLTP□□□□□□□□□□fileio□oltp□□□□□□□□□□□□□□sysbench□□□□□□□□□□□

```
[root@xen-server□]#sysbench--test=fileio help

sysbench 0.4.10:multi-threaded system evaluation benchmark

fileio options:

--file-num=N number of files to create[128]

--file-block-size=N block size to use in all IO operations[16384]
```

```
--file-total-size=SIZE total size of files to create[2G]

--file-test-mode=STRING test mode{seqwr,seqrewr,seqrd,rndrd,rndwr,rndrw}

--file-io-mode=STRING file operations mode{sync,async,fastmmap,slowmmap}[sync]

--file-extra-flags=STRING additional flags to use on opening files{sync,dsync,direct}[]

--file-fsync-freq=N do fsync()after this number of requests(0-don't use fsync())[100]

--file-fsync-all=[on|off]do fsync()after each write operation[off]

--file-fsync-end=[on|off]do fsync()at the end of test[on]

--file-fsync-mode=STRING which method to use for synchronization{fsync,fdatasync}[fsync]

--file-merged-requests=N merge at most this number of IO requests if possible(0-don't merge)[0]

--file-rw-ratio=N reads/writes ratio for combined test[1.5]
```

[illegible]

❑--file-num[illegible]128[illegible]

❑--file-block-size[illegible]InnoDB[illegible]16384[illegible]InnoDB[illegible]16384[illegible]

❑--file-total-size[illegible]2GB[illegible]

❑--file-test-mode[illegible],[illegible]seqwr[illegible]seqrewr[illegible]seqrd[illegible]rndrd[illegible]rndwr[illegible]rndrw[illegible]

❑--file-io-mode[illegible]MMAP[illegible]map[illegible]

❑--file-extra-flags[illegible]API[illegible]

❑--file-fsync-freq[illegible]fsync[illegible]fsync[illegible]

❑--file-fsync-all[illegible]fsync[illegible]off[illegible]

❑--file-fsync-end[illegible]fsync[illegible]on[illegible]

❑--file-fsync-mode[illegible]API[illegible]fdatasync[illegible]fdatasync[illegible]fsync[illegible]

❑--file-rw-ratio[illegible]2:1[illegible]

sysbench[illegible]fileio[illegible]prepare[illegible]run[illegible]cleanup[illegible]prepare[illegible]run[illegible]cleanup[illegible]16[illegible]2GB[illegible]fileio[illegible]

```
[root@xen-server ssd]#sysbench--test=fileio--file-num=16--file-total-size=2G prepare

sysbench 0.4.10:multi-threaded system evaluation benchmark

16 files,131072Kb each,2048Mb total

Creating files for the test...
```

[illegible]16[illegible]2GB[illegible]128MB[illegible]

```
[root@xen-server ssd]#ls-lh

total 2G

-rw-------1 root root 128M Aug 12 10:42 test_file.0

-rw-------1 root root 128M Aug 12 10:42 test_file.1

-rw-------1 root root 128M Aug 12 10:42 test_file.10

-rw-------1 root root 128M Aug 12 10:42 test_file.11

-rw-------1 root root 128M Aug 12 10:42 test_file.12

-rw-------1 root root 128M Aug 12 10:42 test_file.13

-rw-------1 root root 128M Aug 12 10:42 test_file.14

-rw-------1 root root 128M Aug 12 10:42 test_file.15

-rw-------1 root root 128M Aug 12 10:42 test_file.2

-rw-------1 root root 128M Aug 12 10:42 test_file.3

-rw-------1 root root 128M Aug 12 10:42 test_file.4

-rw-------1 root root 128M Aug 12 10:42 test_file.5

-rw-------1 root root 128M Aug 12 10:42 test_file.6
```

```
-rw-------1 root root 128M Aug 12 10:42 test_file.7

-rw-------1 root root 128M Aug 12 10:42 test_file.8

-rw-------1 root root 128M Aug 12 10:42 test_file.9
```

[illegible]16[illegible]

```
[root@xen-server ssd]#sysbench--test=fileio--file-total-size=2G--file-test-mode=rndrd--max-time=180--max-requests=100000000--num-threads=16--init-rng=on--file-num=16--file-extra-flags=direct--file-fsync-freq=0--file-block-size=16384 run
```

[illegible]100 000 000[illegible]180[illegible]

```
[root@xen-server ssd]#sysbench--test=fileio--file-total-size=2G--file-test-mode=rndrd--max-time=180--max-requests=100000000--num-threads=16--init-rng=on--file-num=16--file-extra-flags=direct--file-fsync-freq=0--file-block-size=16384 run

sysbench 0.4.10:multi-threaded system evaluation benchmark

Running the test with following options:

Number of threads:16

Initializing random number generator from timer.

Extra file open flags:16384

16 files,128Mb each

2Gb total file size

Block size 16Kb

Number of random requests for random IO:100000000

Read/Write ratio for combined random IO test:1.50

Calling fsync()at the end of test,Enabled.

Using synchronous I/O mode

Doing random read test

Threads started!

Time limit exceeded,exiting...

(last message repeated 15 times)

Done.
```

```
Operations performed:619908 Read,0 Write,0 Other=619908 Total
Read 9.459Gb Written 0b Total transferred 9.459Gb(53.81Mb/sec)
3443.85 Requests/sec executed
Test execution summary:
total time:180.0044s
total number of events:619908
total time taken by event execution:2878.0750
per-request statistics:
min:0.42ms
avg:4.64ms
max:27.30ms
approx.95 percentile:8.13ms
Threads fairness:
events(avg/stddev):38744.2500/102.69
execution time(avg/stddev):179.8797/0.00
```

[illegible]53.81MB/s[illegible]IOPS[illegible]3443.85[illegible]

[illegible]cleanup[illegible]

```
[root@xen-server ssd]#sysbench--test=fileio--file-num=16--file-total-size=2G cleanup
sysbench 0.4.10:multi-threaded system evaluation benchmark
Removing test files...
```

[illegible]

```
#!/bin/sh
set-u
```

```
set-x

set-e

for size in 8G 64G;do

for mode in seqrd seqrw rndrd rndwr rndrw;do

for blksize in 4096 16384;do

sysbench--test=fileio--file-num=64--file-total-size=$size prepare

for threads in 1 4 8 16 32;do

echo"======testing$blksize in$threads threads"

echo PARAMS$size$mode$threads$blksize[illegible]sysbench-size-$size-mode-$mode-threads-$threads-blksz-$blksize

for i in 1 2 3;do

sysbench--test=fileio--file-total-size=$size--file-test-mode=$mode\

--max-time=180--max-requests=100000000--num-threads=$threads--init-rng=on\

--file-num=64--file-extra-flags=direct--file-fsync-freq=0--file-block-size=$blksize run\

|tee-a sysbench-size-$size-mode-$mode-threads-$threads-blksz-$blksize 2[illegible]1

done

done

sysbench--test=fileio--file-total-size=$size cleanup

done

done

done
```

[illegible]MySQL[illegible]OLTP[illegible]fileio[illegible]prepare[illegible]run[illegible]cleanup[illegible]prepare[illegible]sbtest[illegible]sbtest[illegible]sysbench[illegible]InnoDB[illegible]8000W[illegible]

```
[root@xen-server[illegible]]#sysbench--test=oltp--oltp-table-size=80000000--db-driver=mysql--mysql-socket=/tmp/mysql.sock--mysql-user=root prepare

sysbench 0.4.10:multi-threaded system evaluation benchmark

Creating table'sbtest'...

Creating 80000000 records in table'sbtest'...
```

[illegible]oltp[illegible]

```
sysbench--test=oltp--oltp-table-size=80000000--oltp-read-only=off--init-rng=on--num-threads=16--max-requests=0--oltp-dist-type=uniform--max-time=3600--mysql-user=root--mysql-socket=/tmp/mysql.sock--db-driver=mysql run[illegible]res
```

[illegible]res[illegible]res[illegible]

```
sysbench 0.4.10:multi-threaded system evaluation benchmark

WARNING:Preparing of"BEGIN"is unsupported,using emulation

(last message repeated 15 times)

Running the test with following options:

Number of threads:16

Initializing random number generator from timer.

Doing OLTP test.

Running mixed OLTP test

Using Uniform distribution

Using"BEGIN"for starting transactions

Using auto_inc on the id column

Threads started!

Time limit exceeded,exiting...

(last message repeated 15 times)

Done.

OLTP test statistics:

queries performed:

read:6043324

write:2158330

other:863332

total:9064986

transactions:431666(119.90 per sec.)

deadlocks:0(0.00 per sec.)
```

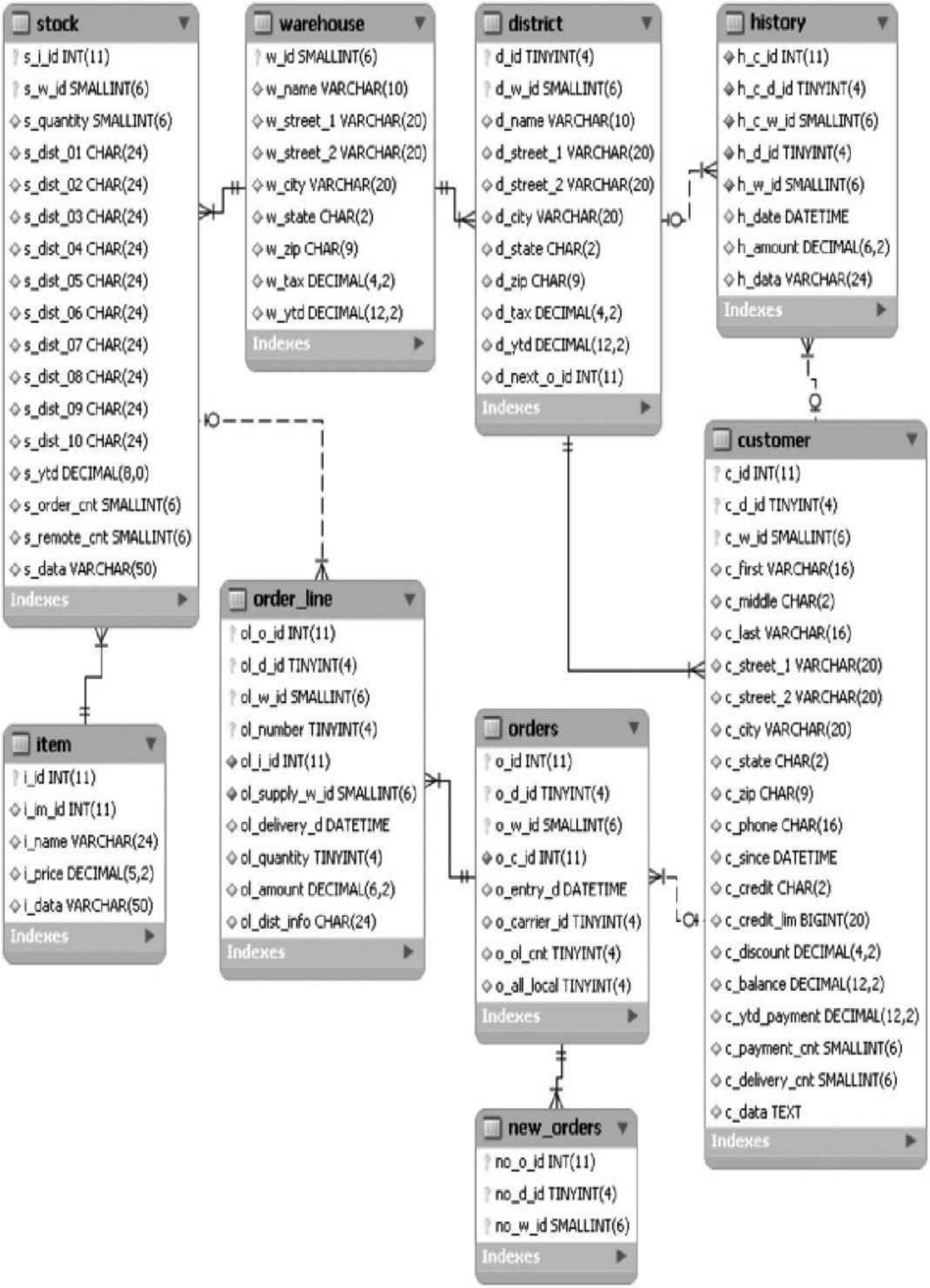
```
read/write requests:8201654(2278.07 per sec.)
other operations:863332(239.80 per sec.)
Test execution summary:
total time:3600.2672s
total number of events:431666
total time taken by event execution:57598.5965
per-request statistics:
min:6.84ms
avg:133.43ms
max:7155.61ms
approx.95 percentile:325.84ms
Threads fairness:
events(avg/stddev):26979.1250/64.14
execution time(avg/stddev):3599.9123/0.06.
```

[illegible]transactions[illegible]TPS[illegible]119.9tps[illegible]sysbench[illegible]OLTP[illegible]TPS[illegible]sysbench[illegible]

9.7.2 mysql-tpcc

TPC[illegible]Transaction Processing Performance Council[illegible]TPC-C[illegible]TPC[illegible]OLTP[illegible]TPC-C[illegible]OLTP[illegible]

TPC-C[illegible]3NF[illegible]10[illegible]3000[illegible]100 000[illegible]9[illegible]9-9[illegible]


stock
s_i_id INT(11)
s_w_id SMALLINT(6)
s_quantity SMALLINT(6)
s_dist_01 CHAR(24)
s_dist_02 CHAR(24)
s_dist_03 CHAR(24)
s_dist_04 CHAR(24)
s_dist_05 CHAR(24)
s_dist_06 CHAR(24)
s_dist_07 CHAR(24)
s_dist_08 CHAR(24)
s_dist_09 CHAR(24)
s_dist_10 CHAR(24)
s_ytd DECIMAL(8,0)
s_order_cnt SMALLINT(6)
s_remote_cnt SMALLINT(6)
s_data VARCHAR(50)
Indexes
warehouse
w_id SMALLINT(6)
w_name VARCHAR(10)
w_street_1 VARCHAR(20)
w_street_2 VARCHAR(20)
w_city VARCHAR(20)
w_state CHAR(2)
w_zip CHAR(9)
w_tax DECIMAL(4,2)
w_ytd DECIMAL(12,2)
Indexes
district
d_id TINYINT(4)
d_w_id SMALLINT(6)
d_name VARCHAR(10)
d_street_1 VARCHAR(20)
d_street_2 VARCHAR(20)
d_city VARCHAR(20)
d_state CHAR(2)
d_zip CHAR(9)
d_tax DECIMAL(4,2)
d_ytd DECIMAL(12,2)
d_next_o_id INT(11)
Indexes
history
h_c_id INT(11)
h_c_d_id TINYINT(4)
h_c_w_id SMALLINT(6)
h_d_id TINYINT(4)
h_w_id SMALLINT(6)
h_date DATETIME
h_amount DECIMAL(6,2)
h_data VARCHAR(24)
Indexes
customer
c_id INT(11)
c_d_id TINYINT(4)
c_w_id SMALLINT(6)
c_first VARCHAR(16)
c_middle CHAR(2)
c_last VARCHAR(16)
c_street_1 VARCHAR(20)
c_street_2 VARCHAR(20)
c_city VARCHAR(20)
c_state CHAR(2)
c_zip CHAR(9)
c_phone CHAR(16)
c_since DATETIME
c_credit CHAR(2)
c_credit_lim BIGINT(20)
c_discount DECIMAL(4,2)
c_balance DECIMAL(12,2)
c_ytd_payment DECIMAL(12,2)
c_payment_cnt SMALLINT(6)
c_delivery_cnt SMALLINT(6)
c_data TEXT
Indexes
order_line
ol_o_id INT(11)
ol_d_id TINYINT(4)
ol_w_id SMALLINT(6)
ol_number TINYINT(4)
ol_i_id INT(11)
ol_supply_w_id SMALLINT(6)
ol_delivery_d DATETIME
ol_quantity TINYINT(4)
ol_amount DECIMAL(6,2)
ol_dist_info CHAR(24)
Indexes
item
i_id INT(11)
i_im_id INT(11)
i_name VARCHAR(24)
i_price DECIMAL(5,2)
i_data VARCHAR(50)
Indexes
orders
o_id INT(11)
o_d_id TINYINT(4)
o_w_id SMALLINT(6)
o_c_id INT(11)
o_entry_d DATETIME
o_carrier_id TINYINT(4)
o_ol_cnt TINYINT(4)
o_all_local TINYINT(4)
Indexes
new_orders
no_o_id INT(11)
no_d_id TINYINT(4)
no_w_id SMALLINT(6)
Indexes

[illegible] 9-9 TPC-C[illegible]

TPC-C[illegible]tpmC[illegible]tpm[illegible]transaction per minute[illegible]C[illegible]TPC[illegible]C[illegible]

tpcc-mysql[illegible]TPC-C[illegible]TPC-C[illegible]https://code.launchpad.net/[illegible]percona-dev/perconatools/tpcc-mysql[illegible]tpcc-mysql[illegible]Linux[illegible]Windows[illegible]http://code.google.com/p/david-mysql-tools/downloads/list[illegible]Windows[illegible]tpcc-mysql[illegible]

tpcc-mysql[illegible]

❑tpcc_load[illegible]9[illegible]

❑tpcc_start[illegible]TPC-C[illegible]

tpcc_load[illegible]

```
[root@xen-server[illegible]]#tpcc_load

*************************************

***###easy###TPC-C Data Loader***

*************************************

usage:tpcc_load[server][DB][user][pass][warehouse]

OR

tpcc_load[server][DB][user][pass][warehouse][part][min_wh][max_wh]

*[part]:1=ITEMS 2=WAREHOUSE 3=CUSTOMER 4=ORDERS
```

[illegible]

❑server[illegible]MySQL[illegible]IP[illegible]

❑DB[illegible]

❑user[illegible]MySQL[illegible]

❏pass□MySQL□□□□

❏warehouse□□□□□□□□□□

□□□tpcc_load□□□□100□□□□□□□tpcc□□□□□□

```
[root@xen-server tpcc-mysql]#mysql tpcc□create_table.sql

[root@xen-server tpcc-mysql]#mysql tpcc□add_fkey_idx.sql

[root@xen-server tpcc-mysql]#tpcc_load 127.0.0.1 tpcc2 root xxxxxx 100

*************************************

***###easy###TPC-C Data Loader***

***************************************

□Parameters□

[server]:127.0.0.1

[DBname]:tpcc2

[user]:root

[pass]:

[warehouse]:100

TPCC Data Load Started...

Loading Item

..................................................5000

..................................................10000

..................................................15000

……□□□

...DATA LOADING COMPLETED SUCCESSFULLY.
```

tpcc_start□□□□□□□□□□

```
[root@xen-server□]#tpcc_start

***************************************

***###easy###TPC-C Load Generator***
```

```
***************************************

usage:tpcc_start[server][DB][user][pass][warehouse][connection][rampup][measure]
```

□□□□□□□□□□

❑connection□□□□□□□□□□

❑rampup□□□□□□□□□□□□□□□□□□□□□□□□□

❑measure□□□□□□□□□□

□□□tpcc_start□□16□□□□□□□□□□□□10□□□□□□□□20□□□□□□

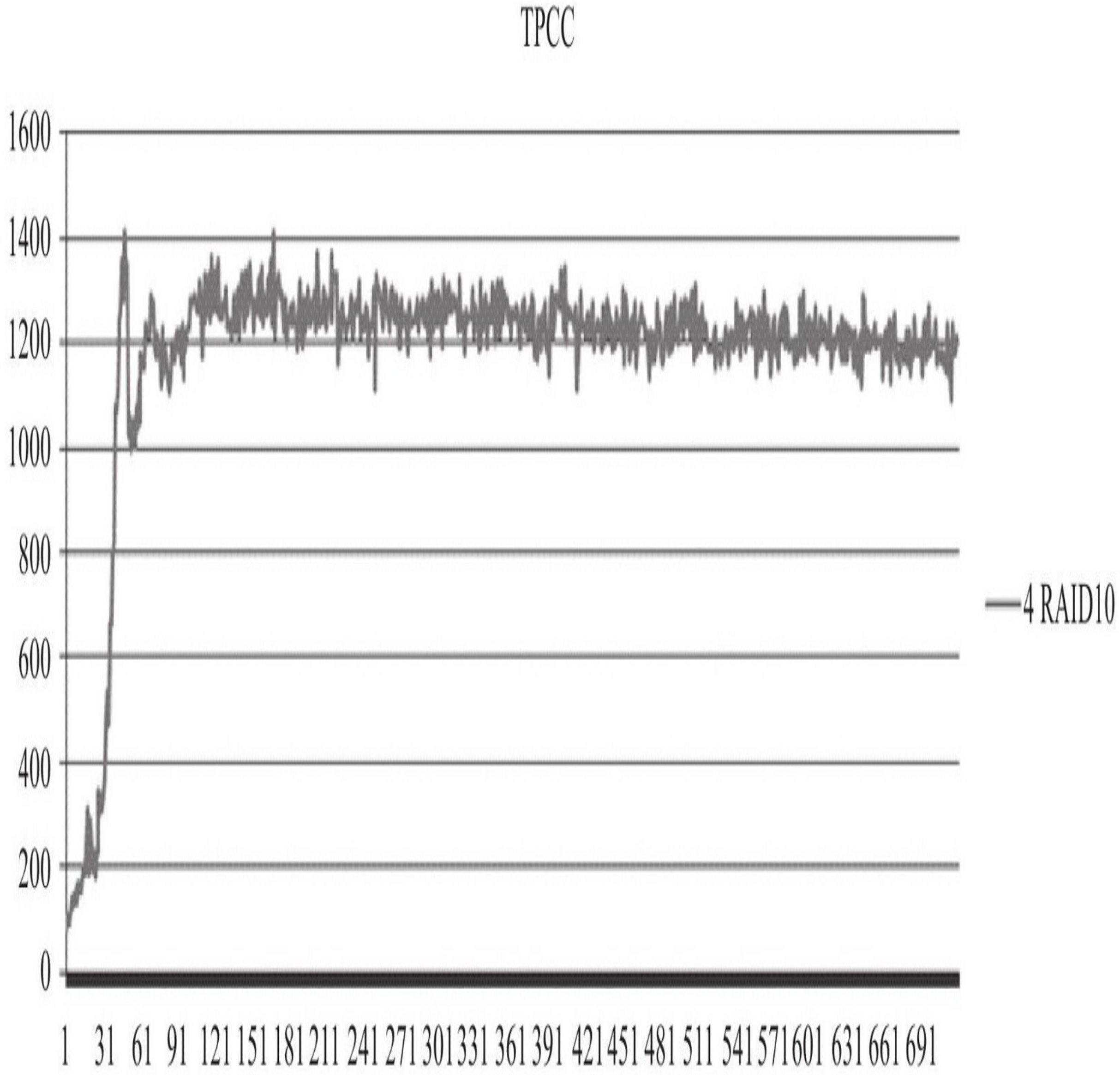
```
[root@xen-server□]#tpcc_start 127.0.0.1 tpcc root xxxxxx 100 16 600 1200

***************************************

***###easy###TPC-C Load Generator***

***************************************

□Parameters□

[server]:127.0.0.1

[DBname]:tpcc

[user]:root

[pass]:xxxxxx

[warehouse]:100

[connection]:16

[rampup]:600(sec.)

[measure]:1200(sec.)

……
```

□□□□□□□□□□□□□□□□□□□□□□□□□

```
RAMP-UP TIME.(1 sec.)

MEASURING START.
```

```
10,624(0):0.4,624(0):0.2,62(0):0.2,63(0):0.6,62(0):0.8

20,990(0):0.2,988(0):0.2,98(0):0.2,99(0):0.4,98(0):0.6

30,1435(0):0.2,1436(0):0.2,144(0):0.2,143(0):0.2,144(0):0.4

40,1736(0):0.2,1739(0):0.2,174(0):0.2,174(0):0.2,174(0):0.4

50,2041(0):0.2,2044(0):0.2,204(0):0.2,204(0):0.2,207(0):0.2

60,2195(0):0.2,2193(0):0.2,220(0):0.2,221(0):0.2,218(0):0.2

70,2332(0):0.2,2335(0):0.2,233(0):0.2,232(0):0.2,234(0):0.2

80,2408(0):0.2,2401(0):0.2,241(0):0.2,239(0):0.2,241(0):0.2

90,2473(0):0.2,2476(0):0.2,247(0):0.2,250(0):0.2,248(0):0.2

100,2350(0):0.2,2347(0):0.2,235(0):0.2,233(0):0.2,235(0):0.2

……
```

□□□□□□10□TPC-C□□□□□□□□TPC-C□□□□□□5□□□□□□□New Order□Payment□Order-Status□Delivery□Stock-Level□□□□□□□New Order□□□□TPC-C□□□□□□□□□□□□□New Order Per 10 Second□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□InnoDB□□□□□10□□□□□□□□□9-10□□□



□ 9-10 New Order Per 10 Second

□tpcc_load□□□□□□□tpmC□□□□New Order Per 10 Second□□□□□□□□New Order Per 10 Second□□□□□□□□6□□□□□□□□tpmC□

```
……

□Constraint Check□(all must be[OK])
```

```
[transaction percentage]

Payment:43.48%(□=43.0%)[OK]

Order-Status:4.35%(□=4.0%)[OK]

Delivery:4.35%(□=4.0%)[OK]

Stock-Level:4.35%(□=4.0%)[OK]

[response time(at least 90%passed)]

New-Order:99.72%[OK]

Payment:99.95%[OK]

Order-Status:99.93%[OK]

Delivery:100.00%[OK]

Stock-Level:100.00%[OK]

□TpmC□

7949.942 TpmC
```

9.8 □□

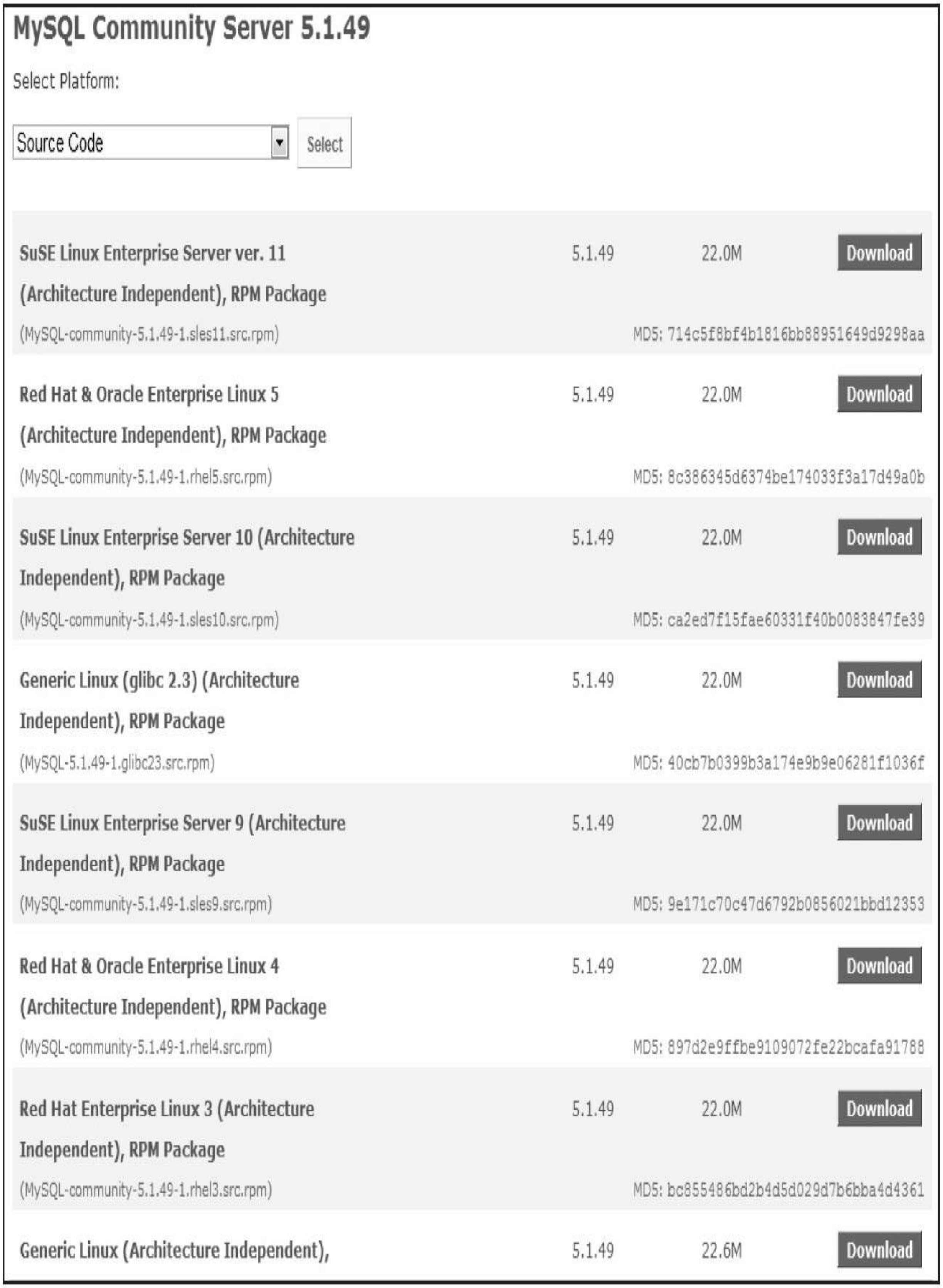
□□□□□□□□□InnoDB□□□□□□□□□□CPU□□□□□□□□□□□□RAID□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Linux□□□□□□□□□□□□□□□□sysbench□tpcc-mysql□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□MySQL□□□□□□□□□□□□□

[illegible]10[illegible] InnoDB[illegible]

InnoDB[illegible]Why[illegible] What[illegible]InnoDB[illegible]

10.1 [illegible]InnoDB[illegible]

InnoDB[illegible]MySQL[illegible]MySQL[illegible][1][illegible]MySQL[illegible]10-1[illegible]



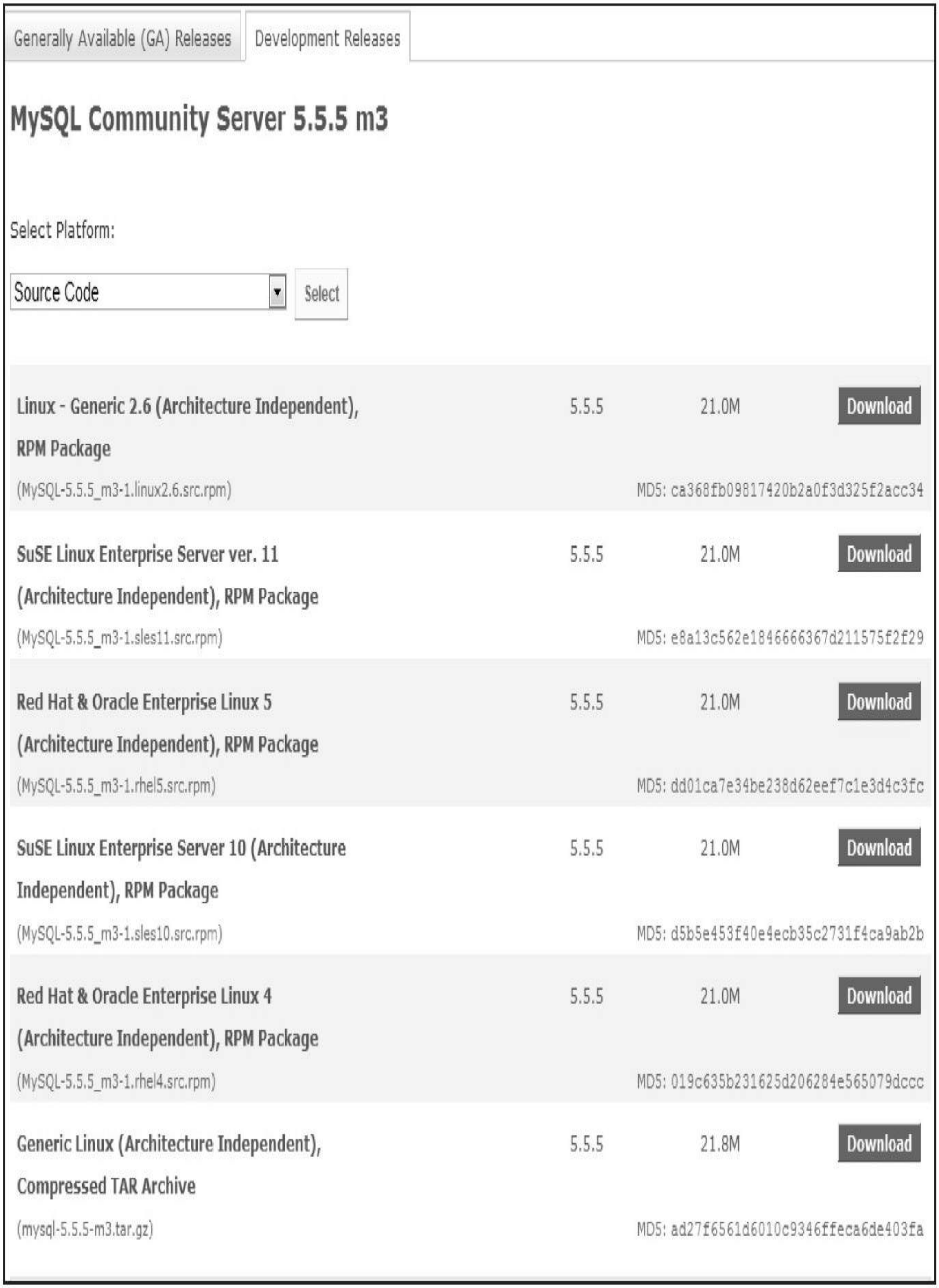
MySQL Community Server 5.1.49

Select Platform:

SuSE Linux Enterprise Server ver. 11 (Architecture Independent), RPM Package (MySQL-community-5.1.49-1.sles11.src.rpm)	5.1.49	22.0M MD5: 714c5f8bf4b1816bb88951649d9298aa	Download
Red Hat & Oracle Enterprise Linux 5 (Architecture Independent), RPM Package (MySQL-community-5.1.49-1.rhel5.src.rpm)	5.1.49	22.0M MD5: 8c386345d6374be174033f3a17d49a0b	Download
SuSE Linux Enterprise Server 10 (Architecture Independent), RPM Package (MySQL-community-5.1.49-1.sles10.src.rpm)	5.1.49	22.0M MD5: ca2ed7f15fae60331f40b0083847fe39	Download
Generic Linux (glibc 2.3) (Architecture Independent), RPM Package (MySQL-5.1.49-1.glibc23.src.rpm)	5.1.49	22.0M MD5: 40cb7b0399b3a174e9b9e06281f1036f	Download
SuSE Linux Enterprise Server 9 (Architecture Independent), RPM Package (MySQL-community-5.1.49-1.sles9.src.rpm)	5.1.49	22.0M MD5: 9e171c70c47d6792b0856021bbd12353	Download
Red Hat & Oracle Enterprise Linux 4 (Architecture Independent), RPM Package (MySQL-community-5.1.49-1.rhel4.src.rpm)	5.1.49	22.0M MD5: 897d2e9ffbe9109072fe22bcafa91788	Download
Red Hat Enterprise Linux 3 (Architecture Independent), RPM Package (MySQL-community-5.1.49-1.rhel3.src.rpm)	5.1.49	22.0M MD5: bc855486bd2b4d5d029d7b6bba4d4361	Download
Generic Linux (Architecture Independent),	5.1.49	22.6M	Download

[illegible] 10-1 MySQL[illegible]

[illegible]Generic Linux[illegible]MySQL[illegible]Download[illegible]GA[illegible]MySQL[illegible]MySQL 5.5.5[illegible]milestone[illegible]GA[illegible]www[illegible]dev[illegible]http://dev.mysql.com/downloads/mysql[illegible]MySQL[illegible]10-2[illegible]

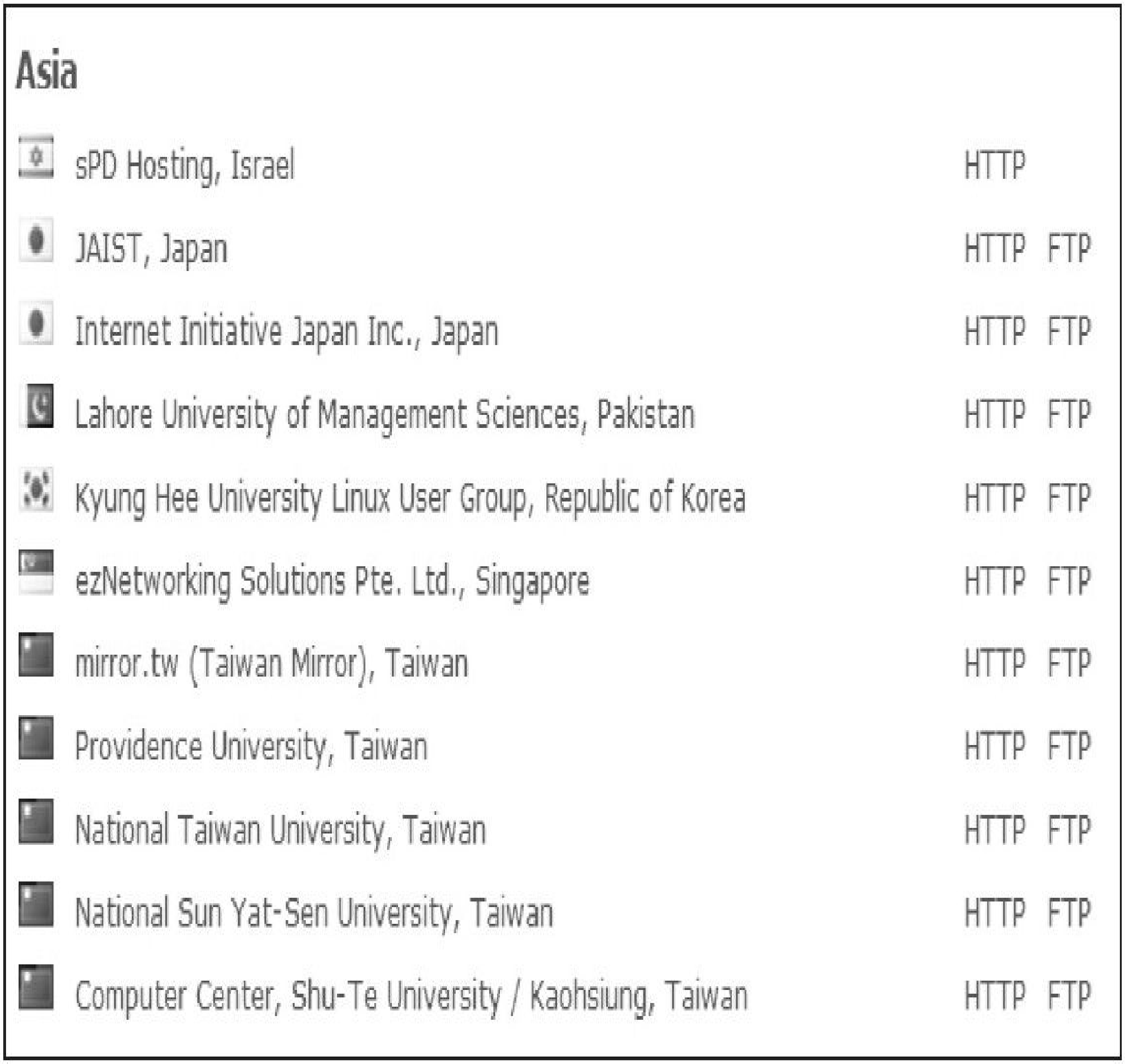


MySQL Community Server 5.5.5 m3

Linux - Generic 2.6 (Architecture Independent), RPM Package (MySQL-5.5.5_m3-1.linux2.6.src.rpm)	5.5.5	21.0M	MD5: ca368fb09817420b2a0f3d325f2acc34
SuSE Linux Enterprise Server ver. 11 (Architecture Independent), RPM Package (MySQL-5.5.5_m3-1.sles11.src.rpm)	5.5.5	21.0M	MD5: e8a13c562e1846666367d211575f2f29
Red Hat & Oracle Enterprise Linux 5 (Architecture Independent), RPM Package (MySQL-5.5.5_m3-1.rhel5.src.rpm)	5.5.5	21.0M	MD5: dd01ca7e34be238d62eef7c1e3d4c3fc
SuSE Linux Enterprise Server 10 (Architecture Independent), RPM Package (MySQL-5.5.5_m3-1.sles10.src.rpm)	5.5.5	21.0M	MD5: d5b5e453f40e4ecb35c2731f4ca9ab2b
Red Hat & Oracle Enterprise Linux 4 (Architecture Independent), RPM Package (MySQL-5.5.5_m3-1.rhel4.src.rpm)	5.5.5	21.0M	MD5: 019c635b231625d206284e565079dccc
Generic Linux (Architecture Independent), Compressed TAR Archive (mysql-5.5.5-m3.tar.gz)	5.5.5	21.8M	MD5: ad27f6561d6010c9346ffeca6de403fa

□ 10-2 MySQL□□□□□□□□□□□

□□“Download”□□□□□□□□□□□□□□□□□□□mysql.com□□□□□□□□□□□MySQL□□“Asia”□□□□□□□□□10-3□□□

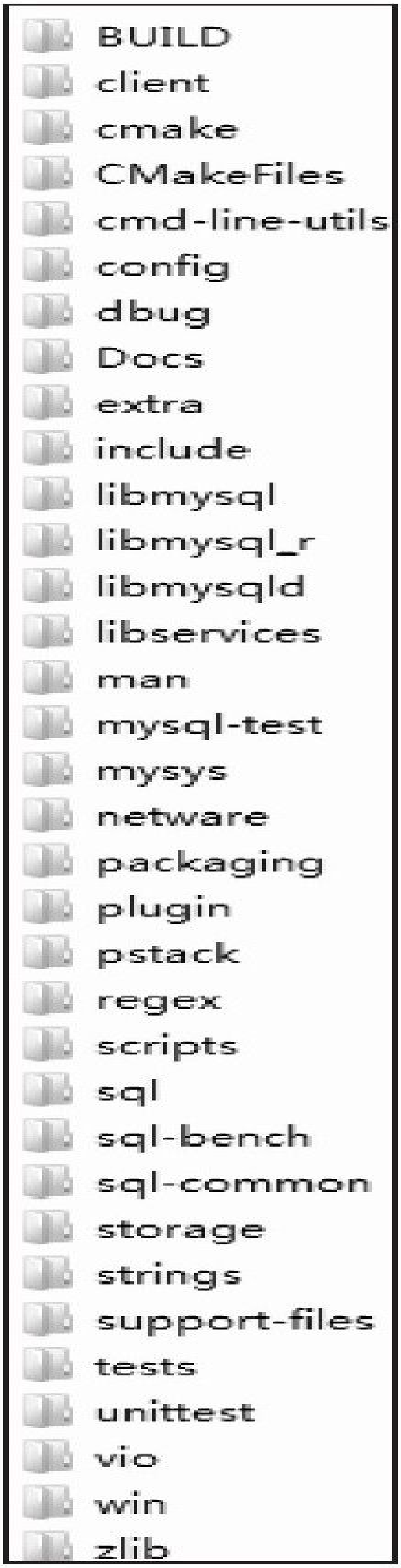


Asia

sPD Hosting, Israel	HTTP	
JAIST, Japan	HTTP	FTP
Internet Initiative Japan Inc., Japan	HTTP	FTP
Lahore University of Management Sciences, Pakistan	HTTP	FTP
Kyung Hee University Linux User Group, Republic of Korea	HTTP	FTP
ezNetworking Solutions Pte. Ltd., Singapore	HTTP	FTP
mirror.tw (Taiwan Mirror), Taiwan	HTTP	FTP
Providence University, Taiwan	HTTP	FTP
National Taiwan University, Taiwan	HTTP	FTP
National Sun Yat-Sen University, Taiwan	HTTP	FTP
Computer Center, Shu-Te University / Kaohsiung, Taiwan	HTTP	FTP

□ 10-3 MySQL□□□□□□

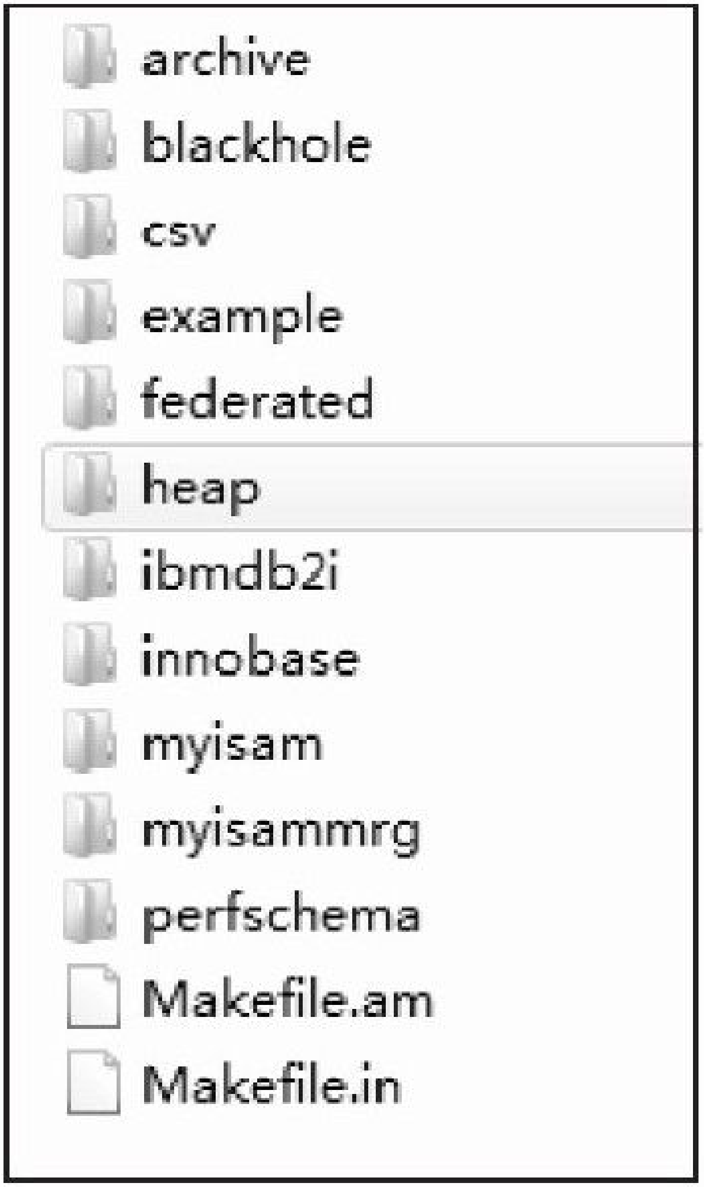
□□□□□□□□tar.gz□□□□□□□□□□Linux□tar□□□Windows□WinRAR□□□□□□□□□□□□□□□□□□□□□□MySQL□□□□□□□□□□□□□□□□□□10-4□



```
BUILD
client
cmake
CMakeFiles
cmd-line-utils
config
dbug
Docs
extra
include
libmysql
libmysql_r
libmysqld
libservices
man
mysql-test
mysys
netware
packaging
plugin
pstack
regex
scripts
sql
sql-bench
sql-common
storage
strings
support-files
tests
unittest
vio
win
zlib
```

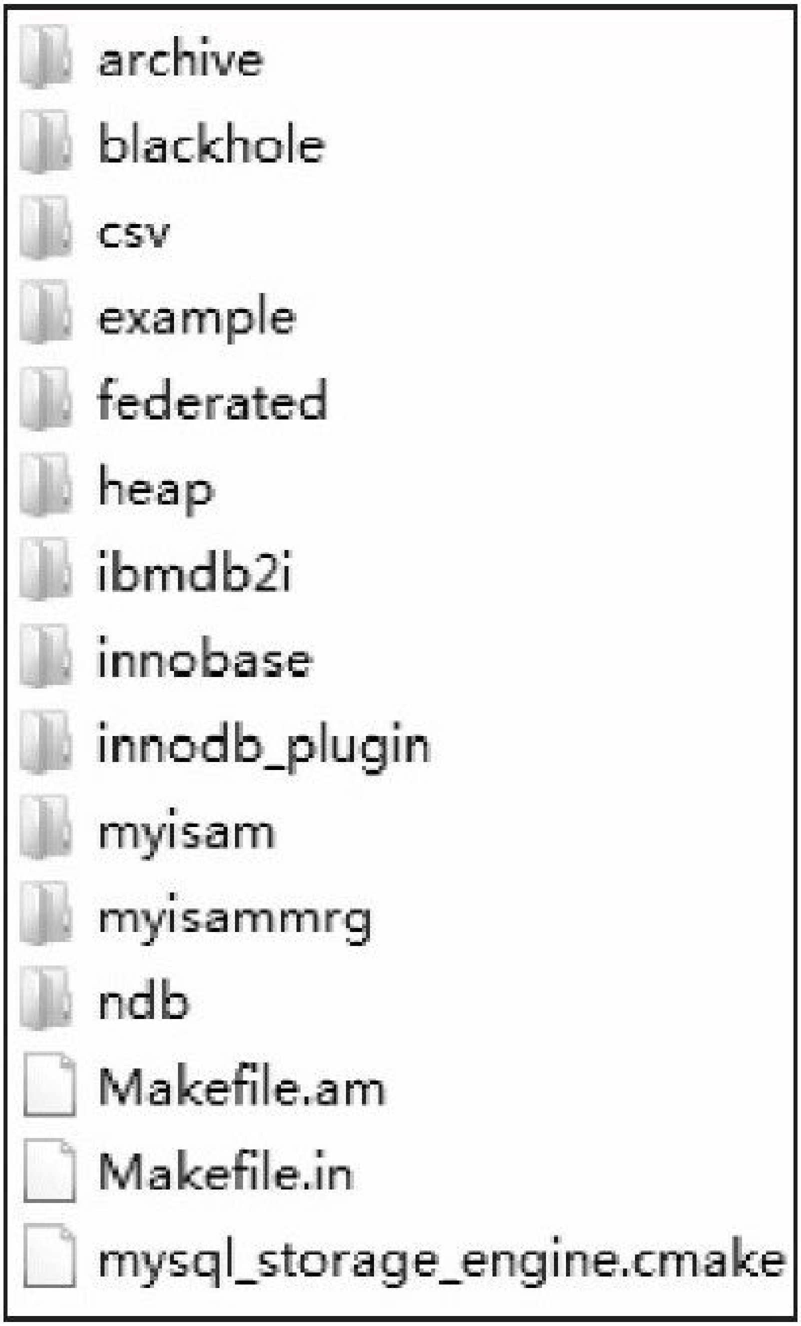
□ 10-4 MySQL□□□□□□□

□□□□□□□□□□□□□□storage□□□□□□□□□□□□□□10-5□□□


archive
blackhole
csv
example
federated
heap
ibmdb2i
innobase
myisam
myisammrg
perfschema
Makefile.am
Makefile.in

□ 10-5 □□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□archive□
blackhole□csv□fedorated□heap□ibmdb2i□myisam□innobase□
□MySQL 5.5□□□□□InnoDB Plugin□□□□□□□InnoDB□□□□□□□□□
MySQL 5.1□□□□□□□□□□□□□□□□□InnoDB□□□□□□□□□□10-6□□□



```
archive
blackhole
csv
example
federated
heap
ibmdb2i
innobase
innodb_plugin
myisam
myisammrg
ndb
Makefile.am
Makefile.in
mysql_storage_engine.cmake
```

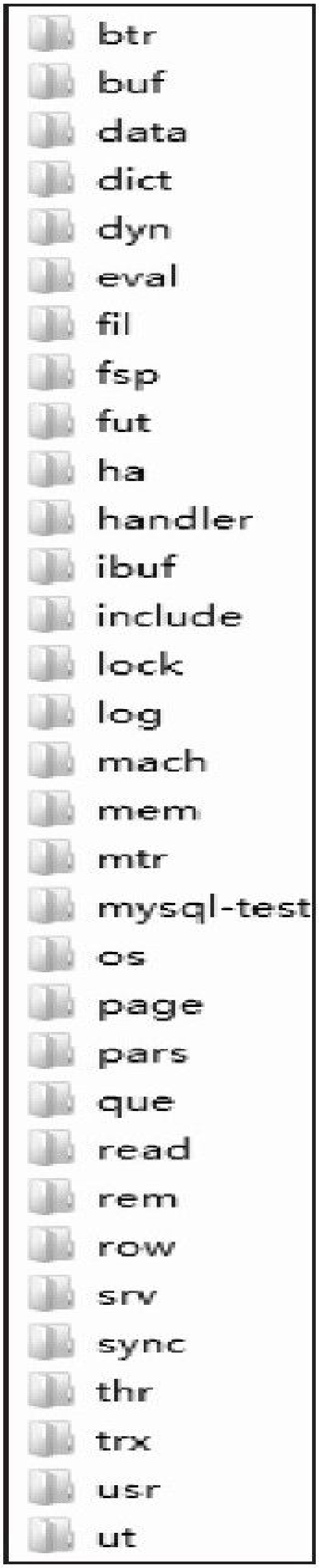
[illegible] 10-6 MySQL 5.1[illegible]

[illegible]innobase[illegible]innodb_plugin[illegible]innobase[illegible]InnoDB [illegible]innodb_plugin[illegible]InnoDB Plugin[illegible] InnoDB Plugin[illegible]MySQL[illegible]innobase[illegible] innodb_plugin[illegible]innobase[illegible]

[1] [illegible]http:‖www.mysql.com/downloads/mysql/[illegible]

10.2 InnoDB[illegible]

[illegible]InnoDB[illegible]10-7[illegible]



btr
buf
data
dict
dyn
eval
fil
fsp
fut
ha
handler
ibuf
include
lock
log
mach
mem
mtr
mysql-test
os
page
pars
que
read
rem
row
srv
sync
thr
trx
usr
ut

[illegible] 10-7 InnoDB[illegible]

❑btr[illegible]B+[illegible]

❑buf[illegible]LRU[illegible]Flush[illegible]

❑dict[illegible]InnoDB[illegible]

❑dyn[illegible]InnoDB[illegible]

❑fil[illegible]InnoDB[illegible]

❑fsp[illegible]file space[illegible]InnoDB[illegible]

❑ha[illegible]

❑handler[illegible]MySQL[illegible]handler[illegible]

❑ibuf[illegible]

❑include[illegible]InnoDB[illegible].h[illegible].ic[illegible]

❑lock[illegible]InnoDB[illegible]S[illegible]X[illegible]

❑log[illegible]

❑mem[illegible]

❑mtr[illegible]

❑os[illegible]

❑page[illegible]

❑row[illegible]

❑srv[illegible]InnoDB[illegible]

❑sync[illegible]InnoDB[illegible]Mutex[illegible]

❑thr[illegible]InnoDB[illegible]

❑trx[illegible]

❑ut[illegible]

10.3 MySQL 5.1□□□□□□□InnoDB□□□

10.3.1 Windows□□□□

□Windows□□□□□□□□Visual Studion 2003□2005□2008□□□□□MySQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

❑CMake□□□□http://www.cmake.org□□□

❑bison□□□□http://gnuwin32.sourceforge.net/packages/bison.htm□□□

□□□□□□□□□configure.js□□□□□□□□□

```
C:\workdir□win\configure.js options
```

options□□□□□□□□□□

❑WITH_INNOBASE_STORAGE_ENGINE□□□InnoDB□□□□□

❑WITH_PARTITION_STORAGE_ENGINE□□□□□□

❑WITH_ARCHIVE_STORAGE_ENGINE□□□Archive□□□□□

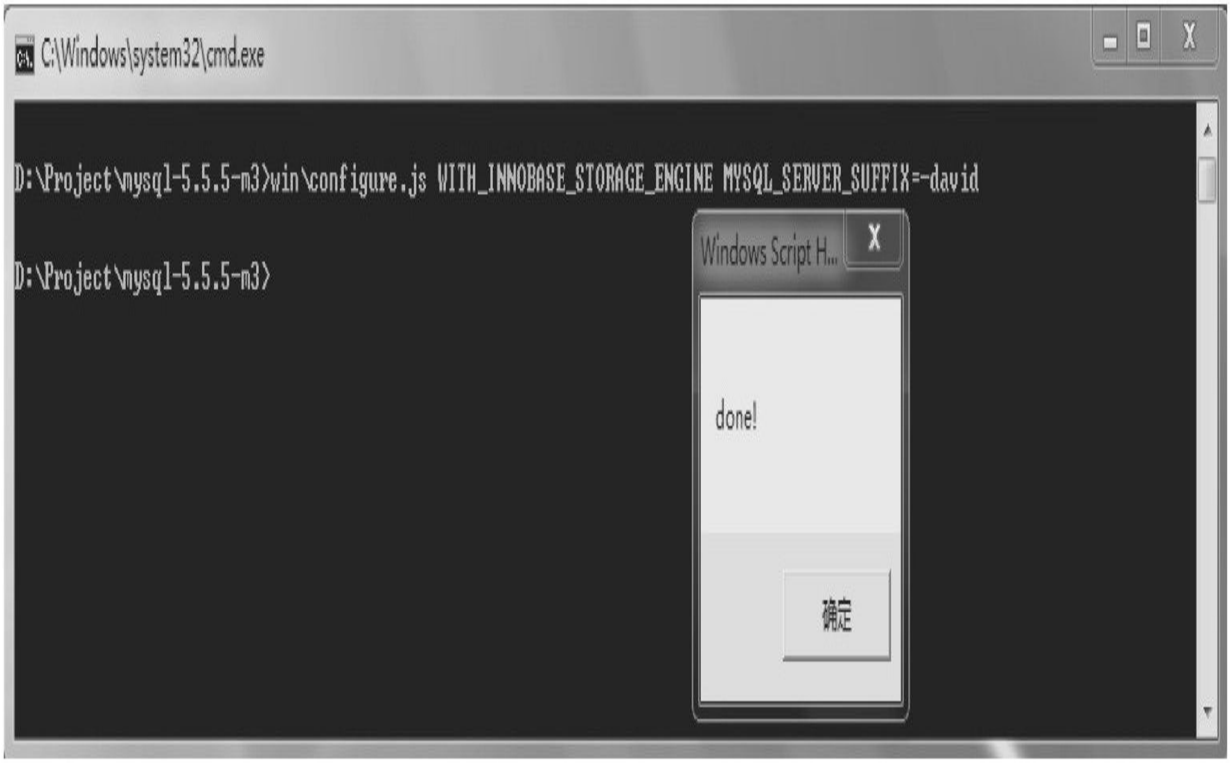
❑WITH_BLACKHOLE_STORAGE_ENGINE□□□Blackhole□□□□□

❑WITH_EXAMPLE_STORAGE_ENGINE□□□Example□□

❑WITH_FEDERATED_STORAGE_ENGINE□□□Federated□□□□□

❑WITH_NDBCLUSTER_STORAGE_ENGINE□□□NDB Cluster□□□□□

□□□□□□□□□□InnoDB□□□□□□□□□□□□□□□□10-8□□□



```
D:\Project\mysql-5.5.5-m3>win\configure.js WITH_INNOBASE_STORAGE_ENGINE MYSQL_SERVER_SUFFIX=-david

D:\Project\mysql-5.5.5-m3>
```

done!

□ 10-8 configure.js□□

□□□□□□□□□□□□Visual Studio 2005□□Visual Studio 2008□win□□□□□build-vsx.bat□□□□□Visual Studio□□□□□□build-vs8.bat□□Visual Studio 2005□build-vs8_x64.bat□□□□□□64□□MySQL□□□□□□□□□□□□32□□□□□□□□□Visual Studio 2008□□□□□□□□□□□□□□□□

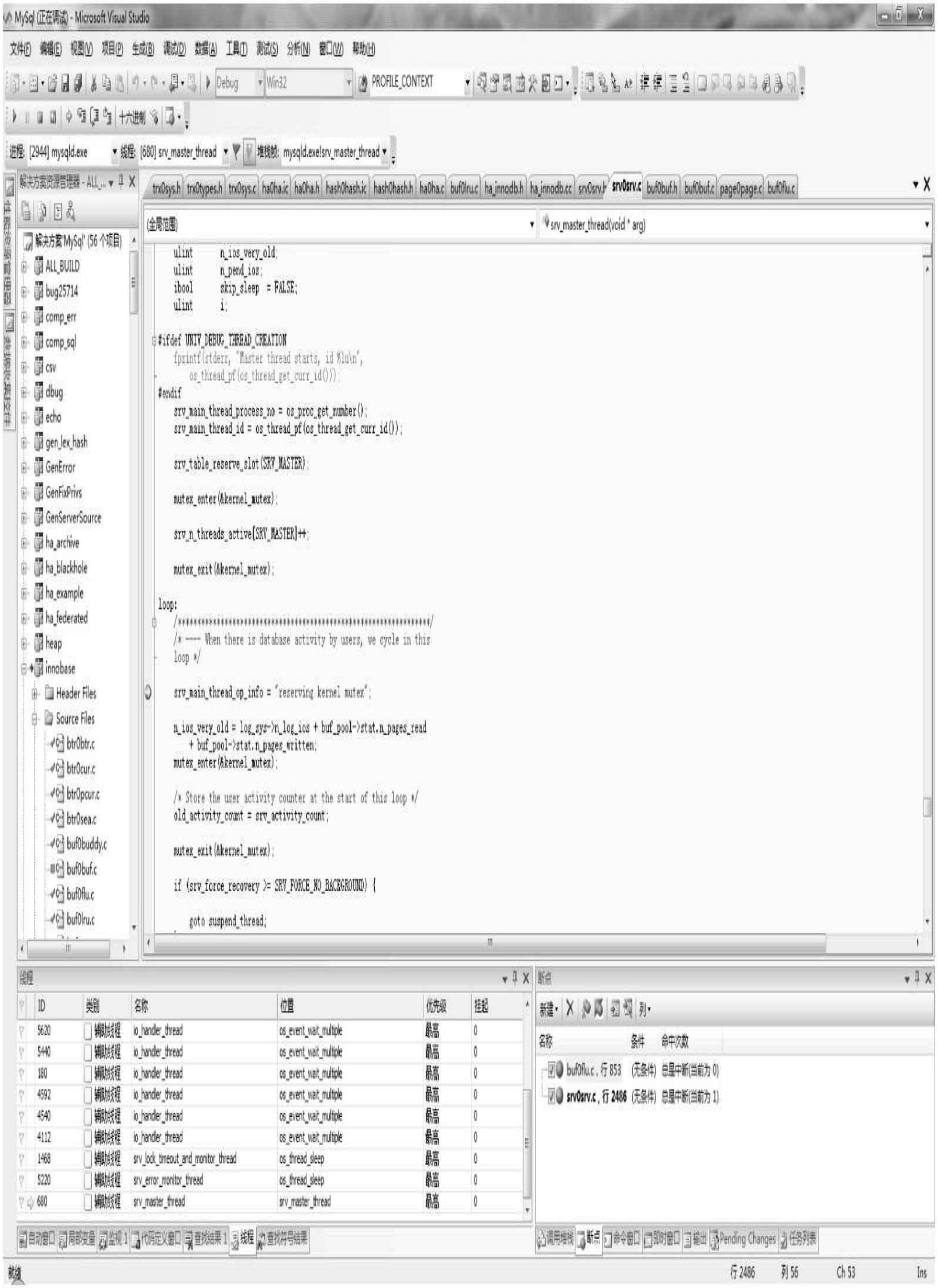
```
D:\Project\mysql-5.5.5-m3□win\build-vs9.bat

--Check for working C compiler:C:/Program Files/Microsoft Visual Studio 9.0/VC/bin/cl.exe

--Check for working C compiler:C:/Program Files/Microsoft Visual Studio 9.0/VC/bin/cl.exe--works

--Detecting C compiler ABI info

--Detecting C compiler ABI info-done

--Check for working CXX compiler:C:/Program Files/Microsoft Visual Studio 9.0/VC/bin/cl.exe

--Check for working CXX compiler:C:/Program Files/Microsoft Visual Studio 9.0/VC/bin/cl.exe--works

--Detecting CXX compiler ABI info

--Detecting CXX compiler ABI info-done

--Check size of void*
```

```
--Check size of void*-done

SIZEOF_VOIDP=4

--Looking for include files HAVE_CXXABI_H

--Looking for include files HAVE_CXXABI_H-not found.

--Looking for include files HAVE_NDIR_H

--Looking for include files HAVE_NDIR_H-not found.

--Looking for include files HAVE_SYS_NDIR_H

--Looking for include files HAVE_SYS_NDIR_H-not found.

--Looking for include files HAVE_ASM_TERMBITS_H

--Looking for include files HAVE_ASM_TERMBITS_H-not found.

--Looking for include files HAVE_TERMBITS_H

--Looking for include files HAVE_TERMBITS_H-not found.

--Looking for include files HAVE_VIS_H

--Looking for include files HAVE_VIS_H-not found.

--Looking for include files HAVE_WCHAR_H

--Looking for include files HAVE_WCHAR_H-found

--Looking for include files HAVE_WCTYPE_H

--Looking for include files HAVE_WCTYPE_H-found

--Looking for include files HAVE_XFS_XFS_H

--Looking for include files HAVE_XFS_XFS_H-not found.

--Looking for include files CMAKE_HAVE_PTHREAD_H

--Looking for include files CMAKE_HAVE_PTHREAD_H-not found.

--Found Threads:TRUE

--Looking for pthread_rwlockattr_setkind_np

--Looking for pthread_rwlockattr_setkind_np-not found

--Performing Test HAVE_SOCKADDR_IN_SIN_LEN

--Performing Test HAVE_SOCKADDR_IN_SIN_LEN-Failed

--Performing Test HAVE_SOCKADDR_IN6_SIN6_LEN

--Performing Test HAVE_SOCKADDR_IN6_SIN6_LEN-Failed

--Cannot find wix 3,installer project will not be generated
```

```
--Configuring done

--Generating done

--Build files have been written to:D:/Project/mysql-5.5.5-m3
```

□□□□□□MySQL.sln□□□□□□□□□□□□□□□□mysqld□□□□□□□□□□□□□□□□□□MySQL□□□□□□□□

□□□□□□□□□□□□□□□□Visual Studio□□□□□□□□□□□□□□□□10-9□□□□InnoDB□□□□□master thread□□□□□



```
	ulint		n_ios_very_old;
	ulint		n_pend_ios;
	ibool		skip_sleep	= FALSE;
	ulint		i;

#ifdef UNIV_DEBUG_THREAD_CREATION
	fprintf(stderr, "Master thread starts, id %lu\n",
		os_thread_pf(os_thread_get_curr_id()));
#endif
	srv_main_thread_process_no = os_proc_get_number();
	srv_main_thread_id = os_thread_pf(os_thread_get_curr_id());

	srv_table_reserve_slot(SRV_MASTER);

	mutex_enter(&kernel_mutex);

	srv_n_threads_active[SRV_MASTER]++;

	mutex_exit(&kernel_mutex);

loop:
	/*****************************************************************/
	/* ---- When there is database activity by users, we cycle in this
	loop */

	srv_main_thread_op_info = "reserving kernel mutex";

	n_ios_very_old = log_sys->n_log_ios + buf_pool->stat.n_pages_read
		+ buf_pool->stat.n_pages_written;
	mutex_enter(&kernel_mutex);

	/* Store the user activity counter at the start of this loop */
	old_activity_count = srv_activity_count;

	mutex_exit(&kernel_mutex);

	if (srv_force_recovery >= SRV_FORCE_NO_BACKGROUND) {

		goto suspend_thread;
```

ID	类别	名称	位置	优先级	挂起
5620	辅助线程	io_handler_thread	os_event_wait_multiple	最高	0
5440	辅助线程	io_handler_thread	os_event_wait_multiple	最高	0
180	辅助线程	io_handler_thread	os_event_wait_multiple	最高	0
4592	辅助线程	io_handler_thread	os_event_wait_multiple	最高	0
4540	辅助线程	io_handler_thread	os_event_wait_multiple	最高	0
4112	辅助线程	io_handler_thread	os_event_wait_multiple	最高	0
1468	辅助线程	srv_lock_timeout_and_monitor_thread	os_thread_sleep	最高	0
5220	辅助线程	srv_error_monitor_thread	os_thread_sleep	最高	0
680	辅助线程	srv_master_thread	srv_master_thread	最高	0

名称	条件	命中次数
☑ buf0flu.c，行 853	(无条件)	总是中断(当前为 0)
☑ srv0srv.c，行 2486	(无条件)	总是中断(当前为 1)

□ 10-9 □□master thread

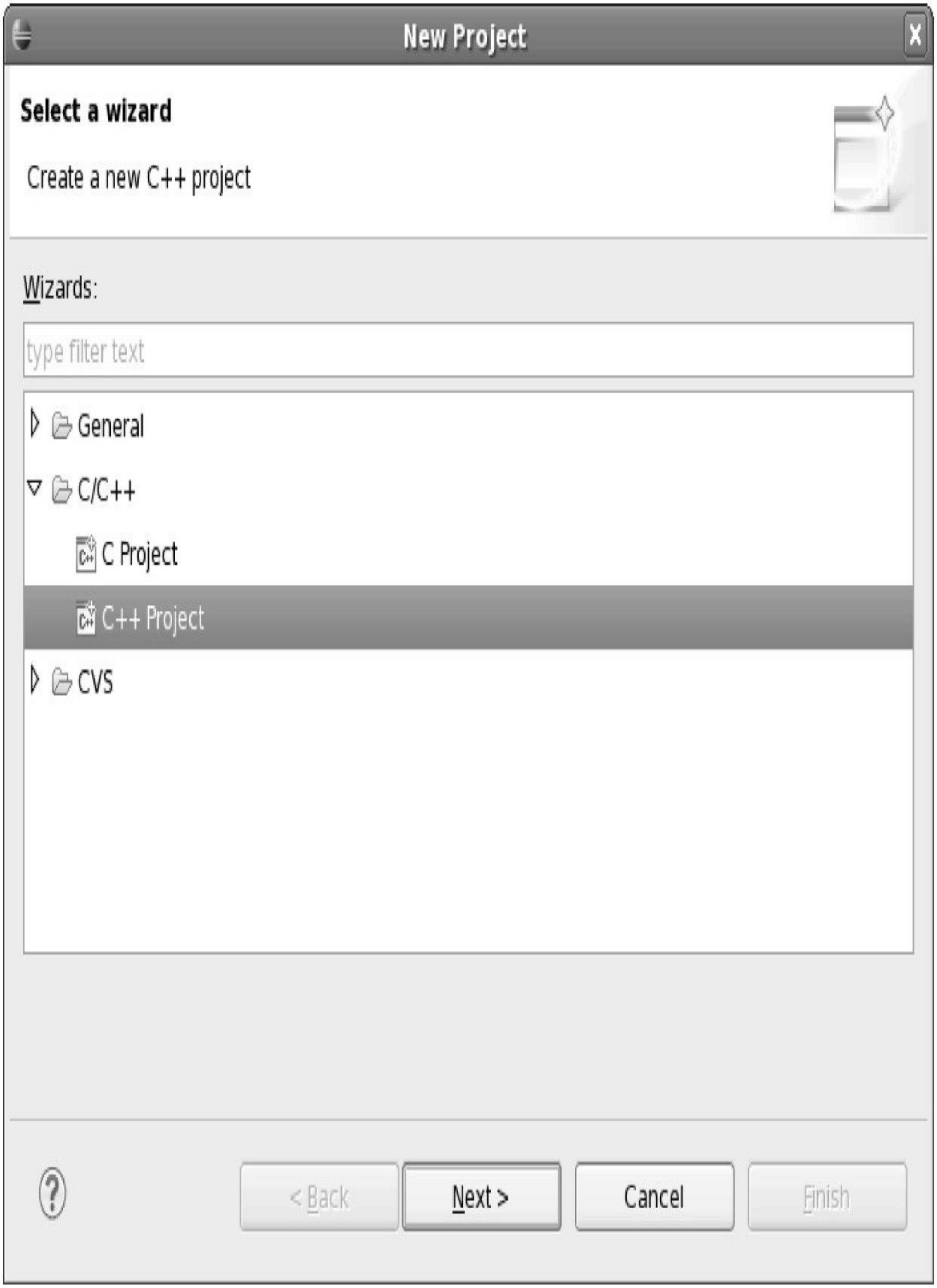
10.3.2 Linux[illegible]

Linux[illegible]Eclipse[illegible]UNIX[illegible]Solaris[illegible]FreeBSD[illegible]MAC[illegible]Eclipse[illegible]http://www.eclipse.org/downloads/[illegible]Eclipse IDE for C/C++Developers[illegible]MySQL[illegible]/root/workspace/mysql-5.5.5-m3[illegible]Make[illegible]Eclipse[illegible]Make[illegible]

```
[root@xen-server mysql-5.5.5-m3]#BUILD/compile-amd64-debug-max-no-ndb-c
```

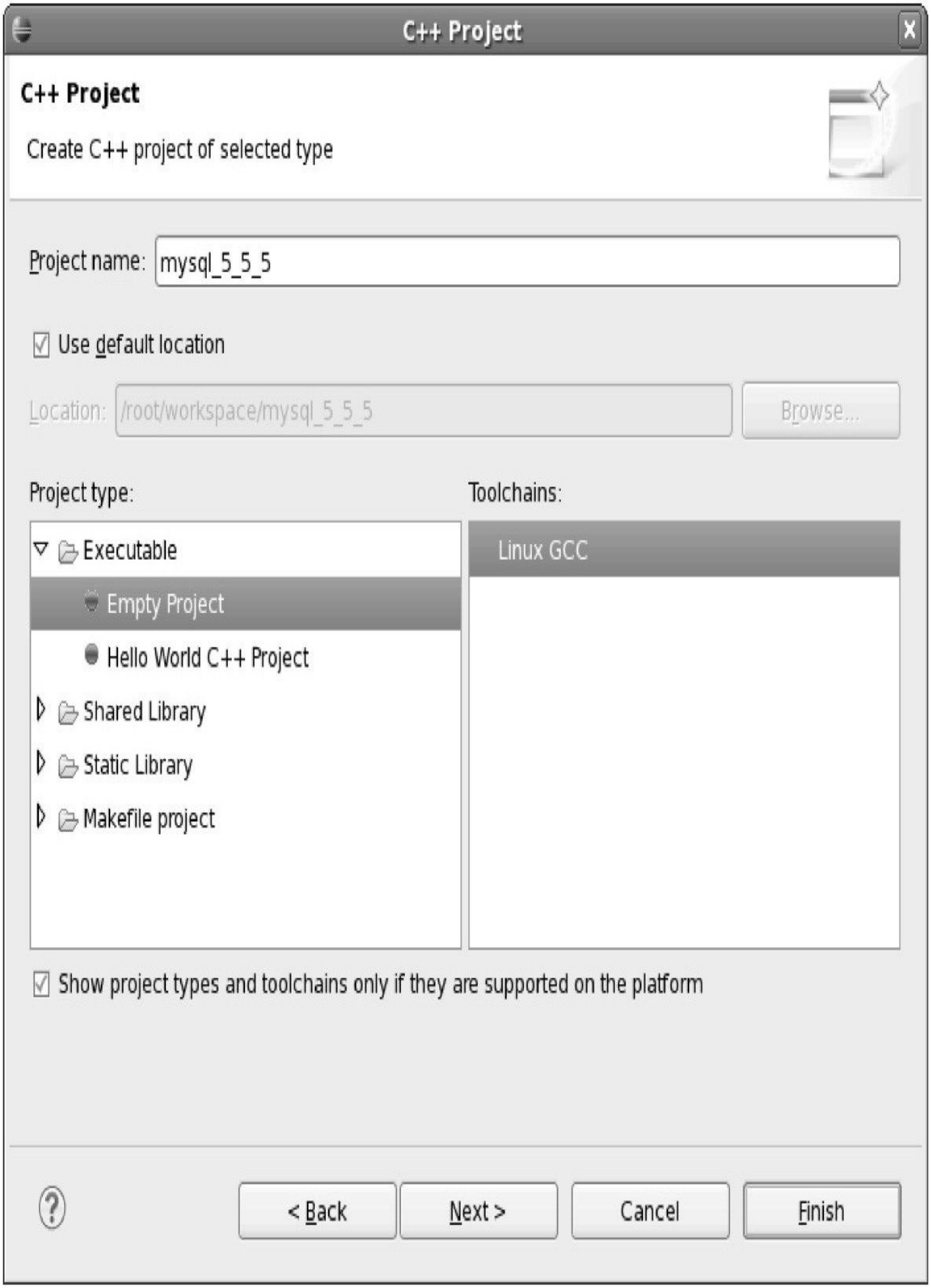
BUILD[illegible]compile[illegible]64[illegible]Linux[illegible]Debug[illegible]compile-amd64-debug-max-no-ndb[illegible]-c[illegible]Make[illegible]

[illegible]Eclipse[illegible]C++[illegible]10-10[illegible]

New Project
Select a wizard
Create a new C++ project
Wizards:
type filter text
General
C/C++
C Project
C++ Project
CVS
< Back
Next >
Cancel
Finish

[illegible] 10-10 [illegible]C++[illegible]

[illegible]mysql_5_5_5[illegible]10-11[illegible]



C++ Project

Create C++ project of selected type

Project name: mysql_5_5_5

☑ Use default location

Location: /root/workspace/mysql_5_5_5

Project type:

- Executable
  - Empty Project
  - Hello World C++ Project
- Shared Library
- Static Library
- Makefile project

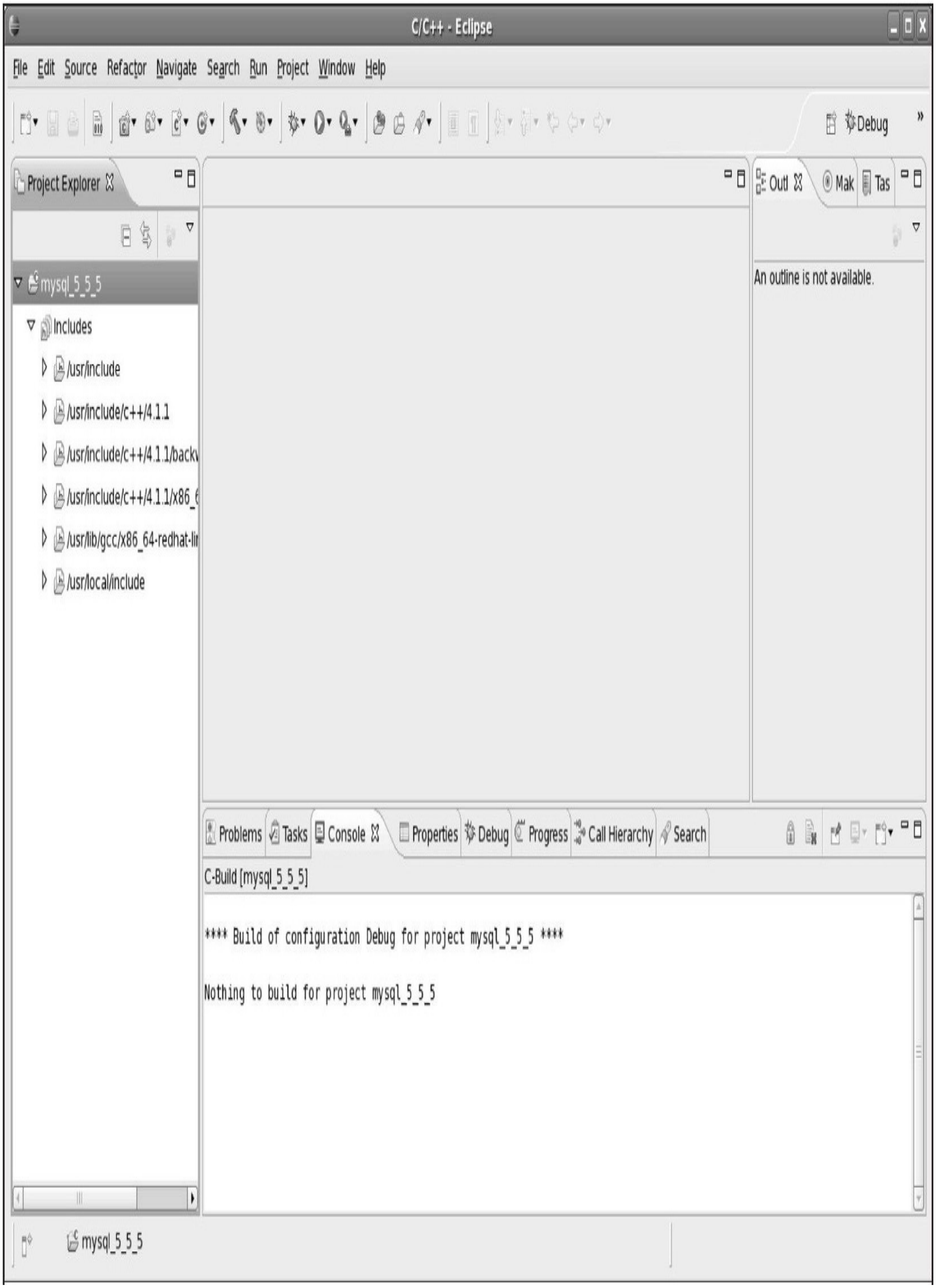
Toolchains:

Linux GCC

☑ Show project types and toolchains only if they are supported on the platform

[illegible] 10-11 [illegible]

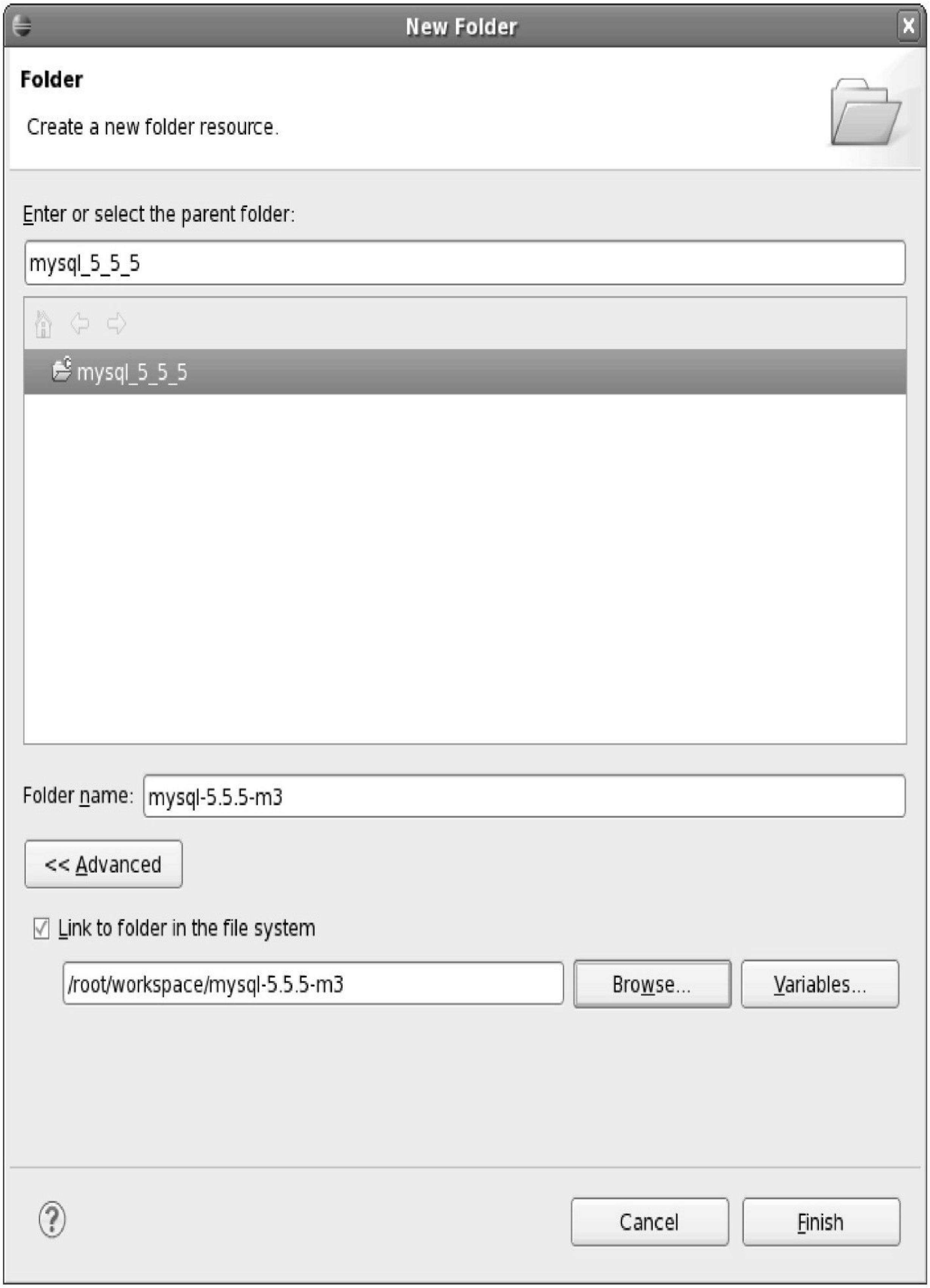
[illegible]Finish[illegible]10-12[illegible]



```
C-Build [mysql_5_5_5]

**** Build of configuration Debug for project mysql_5_5_5 ****

Nothing to build for project mysql_5_5_5
```

[illegible] 10-12 [illegible]C++[illegible]

[illegible]Project Explorer[illegible]mysql_5_5_5[illegible]/root/workspace/mysql-5.5.5-m3[illegible]10-13[illegible]



New Folder

Folder

Create a new folder resource.

Enter or select the parent folder: mysql_5_5_5

mysql_5_5_5

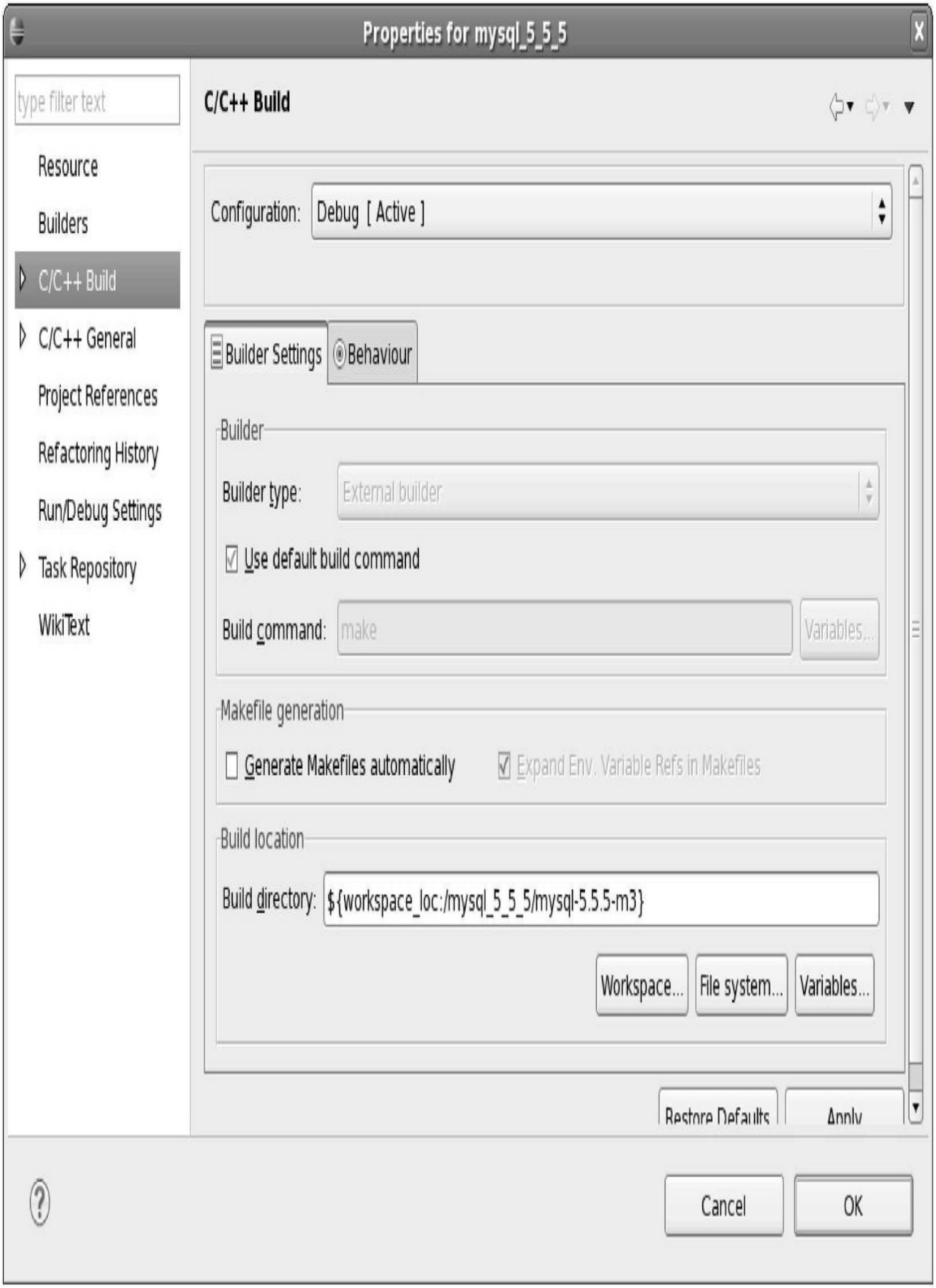
Folder name: mysql-5.5.5-m3

☑ Link to folder in the file system

/root/workspace/mysql-5.5.5-m3

□ 10-13 □□□□□

□□□□□□□□□□□□mysql_5_5_5□□□□□□□□□C/C++Build□□□□□□□□□
□□□Build directory□□□□□□□□□□□□□10-14□□□



Configuration: Debug [Active]

Builder type: External builder

☑ Use default build command

Build command: make

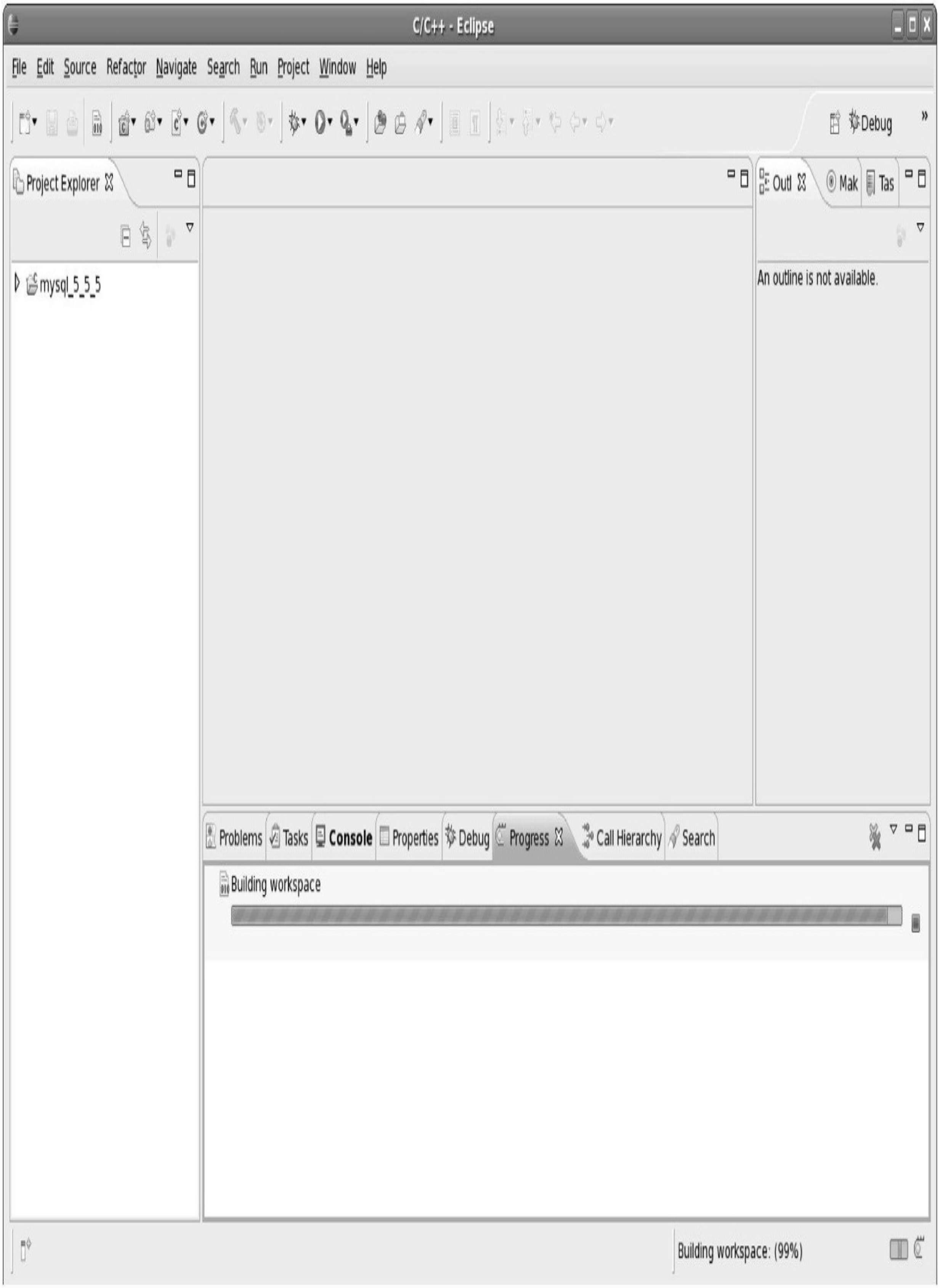
☐ Generate Makefiles automatically

☑ Expand Env. Variable Refs in Makefiles

Build directory: ${workspace_loc:/mysql_5_5_5/mysql-5.5.5-m3}

[illegible] 10-14 [illegible]

[illegible]10-15[illegible]



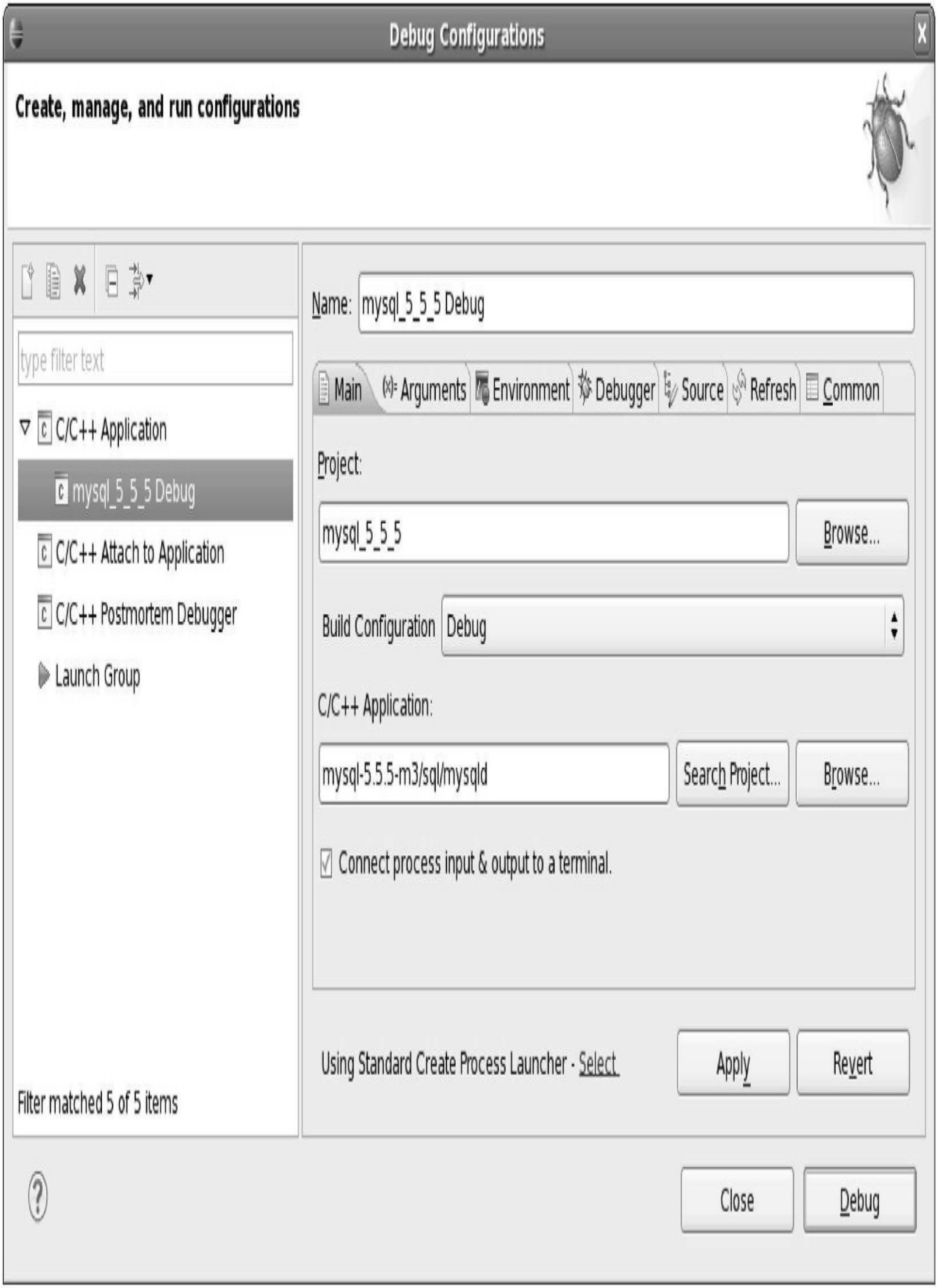
An outline is not available.

Building workspace

Building workspace: (99%)

[illegible] 10-15 [illegible]

[illegible]mysqld[illegible]Debug[illegible]10-16[illegible]



Debug Configurations

Create, manage, and run configurations

Name: mysql_5_5_5 Debug

Project: mysql_5_5_5

Build Configuration: Debug

C/C++ Application: mysql-5.5.5-m3/sql/mysqld

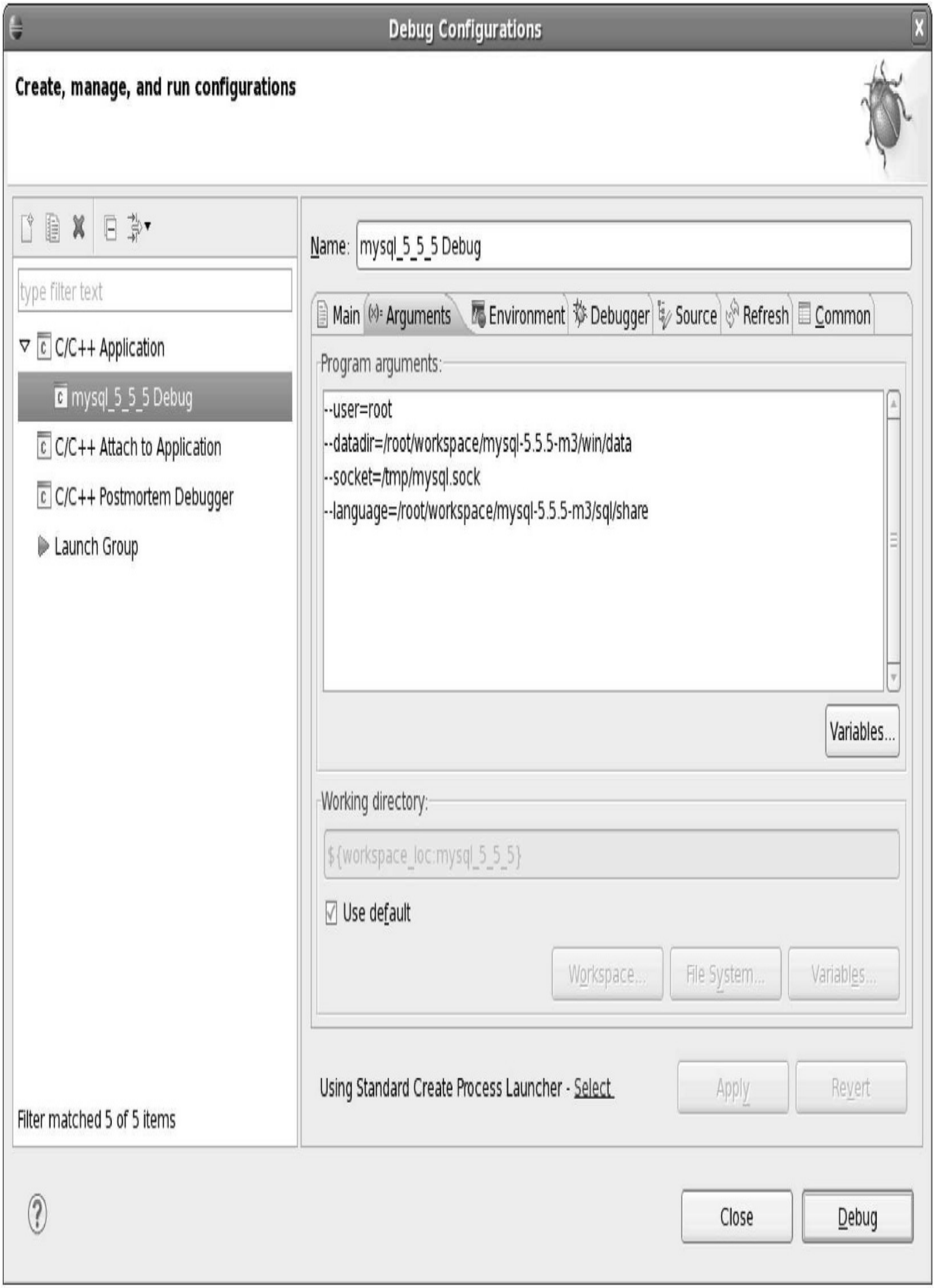
☑ Connect process input & output to a terminal.

Using Standard Create Process Launcher - Select

Filter matched 5 of 5 items

[illegible] 10-16 Debug[illegible]

[illegible]Arguments[illegible]10-17[illegible]



Name: mysql_5_5_5 Debug

Program arguments:

```
--user=root
--datadir=/root/workspace/mysql-5.5.5-m3/win/data
--socket=/tmp/mysql.sock
--language=/root/workspace/mysql-5.5.5-m3/sql/share
```

Working directory:

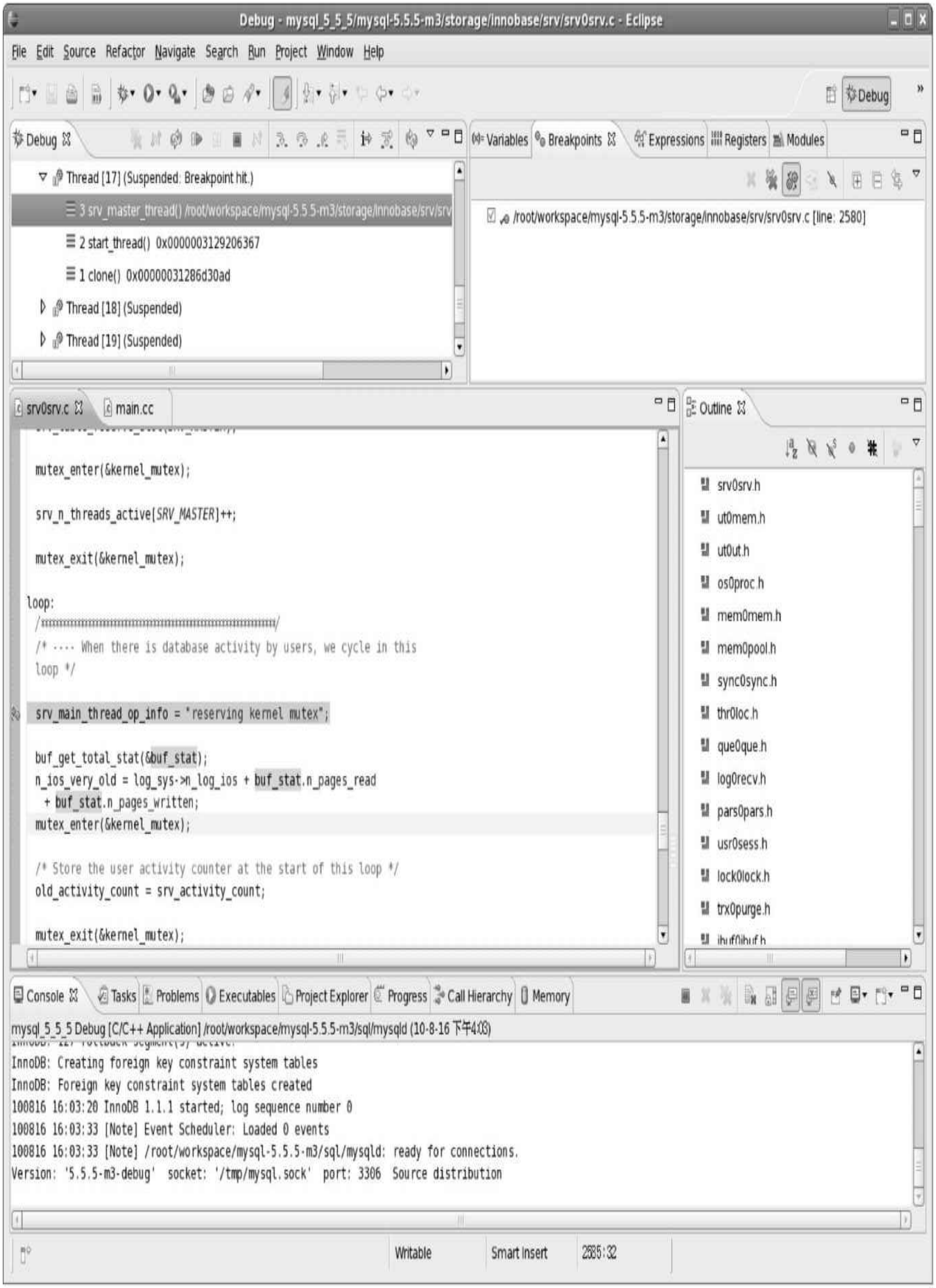
${workspace_loc:mysql_5_5_5}

☑ Use default

Using Standard Create Process Launcher - Select

□ 10-17 □□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□10-18□□□



```
Thread [17] (Suspended: Breakpoint hit.)
    3 srv_master_thread() /root/workspace/mysql-5.5.5-m3/storage/innobase/srv/srv
    2 start_thread() 0x0000003129206367
    1 clone() 0x00000031286d30ad
Thread [18] (Suspended)
Thread [19] (Suspended)
```

```
☑ /root/workspace/mysql-5.5.5-m3/storage/innobase/srv/srv0srv.c [line: 2580]
```

```
	mutex_enter(&kernel_mutex);

	srv_n_threads_active[SRV_MASTER]++;

	mutex_exit(&kernel_mutex);

loop:
	/*****************************************************************/
	/* ---- When there is database activity by users, we cycle in this
	loop */

	srv_main_thread_op_info = "reserving kernel mutex";

	buf_get_total_stat(&buf_stat);
	n_ios_very_old = log_sys->n_log_ios + buf_stat.n_pages_read
		+ buf_stat.n_pages_written;
	mutex_enter(&kernel_mutex);

	/* Store the user activity counter at the start of this loop */
	old_activity_count = srv_activity_count;

	mutex_exit(&kernel_mutex);
```

```
mysql_5_5_5 Debug [C/C++ Application] /root/workspace/mysql-5.5.5-m3/sql/mysqld (10-8-16 下午4:03)
InnoDB: Creating foreign key constraint system tables
InnoDB: Foreign key constraint system tables created
100816 16:03:20 InnoDB 1.1.1 started; log sequence number 0
100816 16:03:33 [Note] Event Scheduler: Loaded 0 events
100816 16:03:33 [Note] /root/workspace/mysql-5.5.5-m3/sql/mysqld: ready for connections.
Version: '5.5.5-m3-debug'  socket: '/tmp/mysql.sock'  port: 3306  Source distribution
```

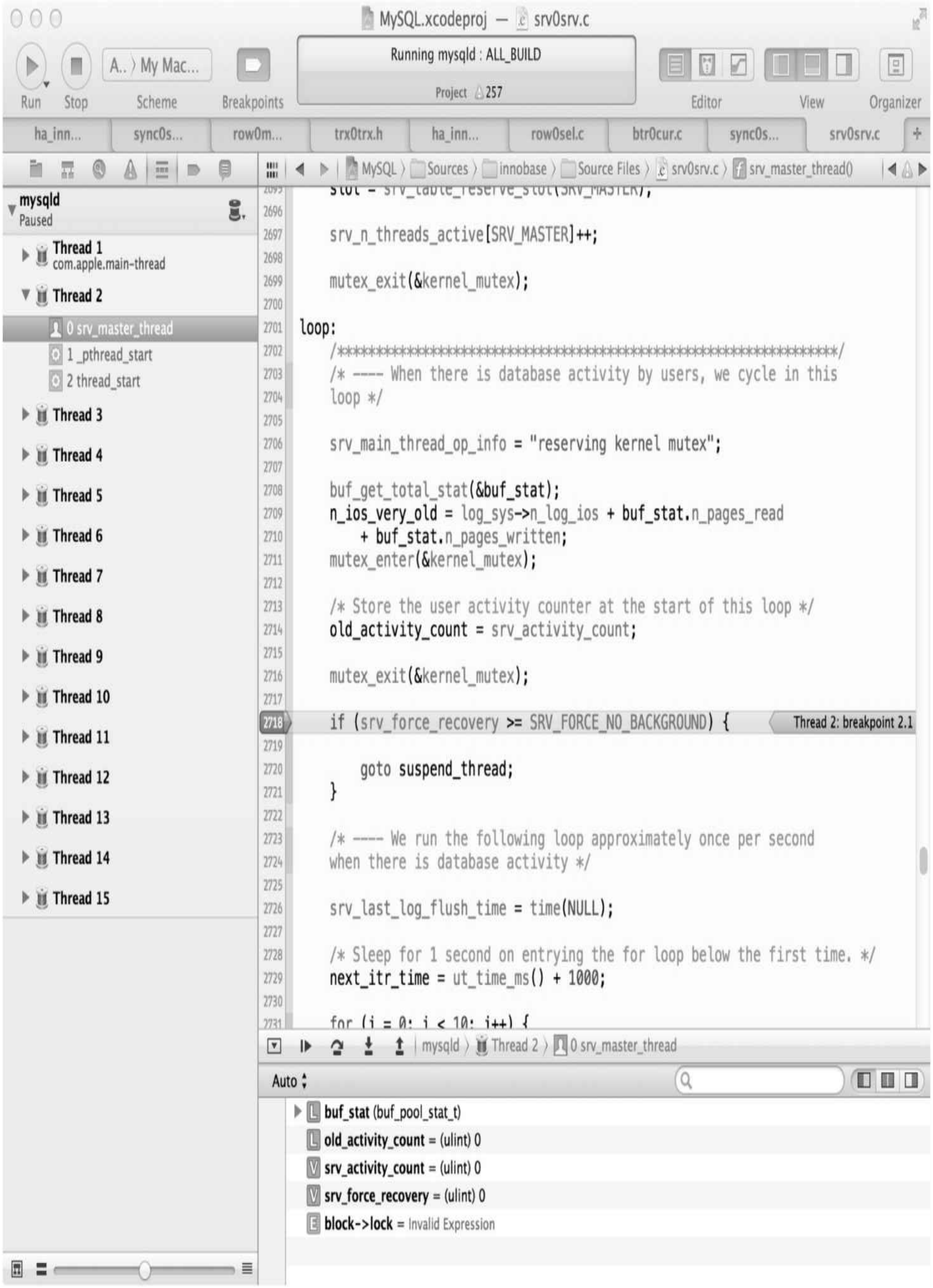
□ 10-18 □Eclipse□□□□

10.4 cmake[illegible]InnoDB[illegible]

MySQL[illegible]5.5[illegible]cmake[illegible]MySQL[illegible]Mac OSX[illegible]xcode[illegible]MySQL[illegible]

```
cd mysql-xxx

mkdir bld

cd bld

cmake..-G Xcode
```

[illegible]MySQL[illegible]bld[illegible]cmake[illegible]xcode[illegible]MySQL.xcodeproj[illegible]MySQL[illegible]InnoDB[illegible]10-19[illegible]



```c
	slot = srv_table_reserve_slot(SRV_MASTER);

	srv_n_threads_active[SRV_MASTER]++;

	mutex_exit(&kernel_mutex);

loop:
	/*****************************************************************/
	/* ---- When there is database activity by users, we cycle in this
	loop */

	srv_main_thread_op_info = "reserving kernel mutex";

	buf_get_total_stat(&buf_stat);
	n_ios_very_old = log_sys->n_log_ios + buf_stat.n_pages_read
		+ buf_stat.n_pages_written;
	mutex_enter(&kernel_mutex);

	/* Store the user activity counter at the start of this loop */
	old_activity_count = srv_activity_count;

	mutex_exit(&kernel_mutex);

	if (srv_force_recovery >= SRV_FORCE_NO_BACKGROUND) {

		goto suspend_thread;
	}

	/* ---- We run the following loop approximately once per second
	when there is database activity */

	srv_last_log_flush_time = time(NULL);

	/* Sleep for 1 second on entrying the for loop below the first time. */
	next_itr_time = ut_time_ms() + 1000;

	for (i = 0; i < 10; i++) {
```

```
▶ L buf_stat (buf_pool_stat_t)
  L old_activity_count = (ulint) 0
  V srv_activity_count = (ulint) 0
  V srv_force_recovery = (ulint) 0
  E block->lock = Invalid Expression
```

[illegible] 10-19 Mac OS X[illegible]xcode[illegible]MySQL[illegible]

[illegible]cmake[illegible]MySQL[illegible]cmake[illegible]MySQL [illegible]

10.5 [illegible]

MySQL[illegible]InnoDB[illegible]Visual Studio[illegible]Eclipse[illegible]Oracle[illegible]Microsoft SQL Server[illegible]DB2[illegible]